

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 30, 2004, 12:45:36 ; Search time 6341 Seconds
(without alignments)
11388.010 Million cell updates/sec

Title: US-09-604-231-1

Perfect score: 1527

Sequence: 1 ctcattggcattcgcgcgtt.....gttgaaaccttgagtgttcg 1527

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4526729 segs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1527	100.0	1527	6	AX069134	AX069134 Sequence
2	1527	100.0	1527	6	AX469820	AX469820 Sequence
3	1515	99.2	320550	1	AP005282	Corynebacterium glutamicum
4	1515	99.2	349136	1	AX927155	Corynebacterium glutamicum
5	1515	99.2	349980	6	AX127152	Sequence
6	1492	97.7	1983	6	AX122988	Sequence
7	1492	97.7	1983	6	BD165105	Novel pol
8	1412.6	92.5	5969	6	BD093238	DNA encod
9	1107.4	72.5	1109	6	AX069136	Sequence
10	308.4	20.2	343050	1	AL935252	Lactobac
11	287.6	18.8	21838	1	PDCAFOPR	Streptococ
12	287.6	18.8	21839	1	PPSRFOP	Streptococ
13	253.4	16.6	2788	1	AF401046	Lactobac
14	250.8	16.4	5800	1	LL297015	Lactobac
15	231.8	15.2	10264	1	AB004395	Lactobac
16	230.8	15.1	2508	1	STRSCRA	Streptococ
17	229.2	15.0	9979	1	AB015011	Streptococ
18	221.6	14.5	4750	1	AY177419	Lactobac
19	218.8	14.3	177911	1	AE017206	Lactobac

20	218.8	14.3	349980	6	AX926712	Sequence
21	218.6	14.3	10085	1	AE006222	Pasteurel
22	215.2	14.1	2955	1	VIBSCRAK	Vibrio algi
23	210.6	13.8	2655	1	SXSCRA	Xylophaga
24	205.4	13.5	4158	1	BSTREAPR	Bacillus s
25	201.8	13.2	22197	1	D83967	Bacillus s
26	201.8	13.2	37900	1	D86417	Bacillus s
27	201.8	13.2	197409	1	BSUR0005	Bacillus s
28	199.6	13.1	110000	1	EX571856_25	Continuation (26 o
29	196.4	12.9	39425	1	AC090968	Staphyloc
30	196.4	12.9	110000	1	EX571857_24	Continuation (25 o
31	196.4	12.9	301250	1	AP004830	Staphyloc
32	194.8	12.8	297850	1	AP003137	Staphyloc
33	194.8	12.8	342350	1	AP003365	Staphyloc
34	190.6	12.5	2913	1	AF269422	Staphyloc
35	190.6	12.5	2913	6	AR485378	Staphyloc
36	190.6	12.5	2913	6	AX144742	Sequence
37	190.6	12.5	300698	1	AE016750	Staphyloc
38	183	12.0	299850	1	AP001514	Bacillus
39	178.4	11.7	5652	1	AY131006	Pseudomon
40	177.8	11.6	975	6	AX432414	Sequence
41	174.6	11.4	861	6	AX621496	Sequence
42	174	11.4	215050	1	AL646084	Ralstonia
43	173.6	11.4	302998	1	AE016921	Chromobac
44	171.6	11.2	1491	6	AR396402	Sequence
45	171.6	11.2	66320	1	AE016833	Enterococ

ALIGNMENTS

RESULT 1	AX069134	AX069134	Sequence 1 from Patent WO0102583.	1527 bp	DNA	linear	PAT 25-JAN-2001
LOCUS	AX069134	AX069134.1	GI:12579016				
DEFINITION	Sequence 1 from Patent WO0102583.						
ACCESSION	AX069134						
VERSION	AX069134.1						
KEYWORDS	Corynebacterium glutamicum						
SOURCE	Corynebacterium glutamicum						
ORGANISM	Corynebacterium glutamicum						
REFERENCE	1						
AUTHORS	Pompejus, M., Kroeger, B., Schroeder, H., Zelder, O. and Haberhauer, G.						
TITLE	Corynebacterium glutamicum genes encoding phosphoenolpyruvate						
JOURNAL	sugar phosphotransferase system proteins						
FEATURES	Location/Qualifiers						
source	1. .1527						
CDS	/organism="Corynebacterium glutamicum"						
	/mol_type="unassigned DNA"						
	/db_xref="taxon:1718"						
	101. .1507						
	/note="unnamed protein product; RXS00315"						
	/codon_start=1						
	/transl_table=11						
	/protein_id="CAC27276.1"						
	/db_xref="GI:12579017"						
	/translation="MAMVFPPLVNGYDVAATMAAGEMPWLSFLGLDVAQAGYQGTVLPL						
	LQGLYDFGPGVGLLFGVSPVITVLGHQSPPELELFNCGGSPFIFATASMANIAQ						
	VIVWSVILATIEKFLKRLKGTADFLITPVLTLTGLTFTAI GPAMRWGVDVLAHG						
	GAACLVAFKASEKLGAGSGVAVLGIETEPALFVGNLRLRPPFFIGTAAIGG						
	ALLALFNKAVNALGAAGLGVVSDAPDMVFLVCVAVTFFAFGAATAGLYLVRRN						
	GSIDPDAAPVPAGTTKAEAPAFNSDSTIIQAPLTGEALSSVSDAMFASGKL						
	GSGLVAVPTKQGVSPVSGKIVPAFNSDSTIIQAPLTGEALSSVSDAMFASGKL						
	THFNELKQGVDEKAGELLCEPDI DAIRKAGVEVTTPIVSNKYKTGTPVNTYGLGIE						
	AGANLLNVAKKEAVPATP"						

ORIGIN

Query Match 100.0%; Score 1527; DB 6; Length 1527;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1527; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 1527; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	CTCATGCACTCGCGCTTCGGCTTCCTCCAGTGTGGTTCGCTTCAACGCAACCAAG	60
Db	1	CTCATGCACTCGCGCTTCGGCTTCCTCCAGTGTGGTTCGCTTCAACGCAACCAAG	60
QY	61	CGTTTCGGCGGCAATGAGTTCTTGGCGCGCGCTATTTGGTATGCGATGGTTCGCGAG	120
Db	61	CGTTTCGGCGGCAATGAGTTCTTGGCGCGCGCTATTTGGTATGCGATGGTTCGCGAG	120
QY	121	CTTGGTGAACGGCTACGAGTGGCGCGCAACATGCTCGCGGCAAAATGCAATGTGCTC	180
Db	121	CTTGGTGAACGGCTACGAGTGGCGCGCAACATGCTCGCGGCAAAATGCAATGTGCTC	180
QY	181	CCTGTTTGGTTTAGATGTTTGCACAGCCGGTTACCAAGGCAACCGTCTTCTGCTGGT	240
Db	181	CCTGTTTGGTTTAGATGTTTGCACAGCCGGTTACCAAGGCAACCGTCTTCTGCTGGT	240
QY	241	GGTTTCTTGGATTTCTGGCAACGATCGAGAGTTCTTGCACACGACTCAAGGGCACTGC	300
Db	241	GGTTTCTTGGATTTCTGGCAACGATCGAGAGTTCTTGCACACGACTCAAGGGCACTGC	300
QY	301	AGACTTCTCGATCACTCCAGTGTGAGTGTCTGCTCACCGGATTCCTTACATTCATCGC	360
Db	301	AGACTTCTCGATCACTCCAGTGTGAGTGTCTGCTCACCGGATTCCTTACATTCATCGC	360
QY	361	CAATTGGCCACGCAATGCGCTGGCGGCGATGTGCTGGCAACCGTCTACAGGCACTTTA	420
Db	361	CAATTGGCCACGCAATGCGCTGGCGGCGATGTGCTGGCAACCGTCTACAGGCACTTTA	420
QY	421	TGATTTCCGTTGCTCAGTCCGCGGTCTGCTCTTGGTCTGCTTACTCAACCAATCGTCAT	480
Db	421	TGATTTCCGTTGCTCAGTCCGCGGTCTGCTCTTGGTCTGCTTACTCAACCAATCGTCAT	480
QY	481	CACTGGTCTGCACCACTCTTCCCGCAATGAGCTGGAGCTGTTTAAACAGGCTGATC	540
Db	481	CACTGGTCTGCACCACTCTTCCCGCAATGAGCTGGAGCTGTTTAAACAGGCTGATC	540
QY	541	CTTCATCTTCGCAACCGGCACTATGCTAATATCGCCAGGGTGGCGCATGTTTGGCAGT	600
Db	541	CTTCATCTTCGCAACCGGCACTATGCTAATATCGCCAGGGTGGCGCATGTTTGGCAGT	600
QY	601	GTTCTTCTGGCGAAGAGTGAAAGCTCAAGGCTTTCGAGGTGCTTTCAGGTGCTCCGC	660
Db	601	GTTCTTCTGGCGAAGAGTGAAAGCTCAAGGCTTTCGAGGTGCTTTCAGGTGCTCCGC	660
QY	661	TGTTCTTGGTATTAACGAGCTCGATCTTCGCTGTGAACTTCGCTCGCGCTGGCGGTT	720
Db	661	TGTTCTTGGTATTAACGAGCTCGATCTTCGCTGTGAACTTCGCTCGCGCTGGCGGTT	720
QY	721	CTTCATCGGTATCGGTACCGCACTATCGGTGGCGCTTTGATGCACTCTTTAATATCAA	780
Db	721	CTTCATCGGTATCGGTACCGCACTATCGGTGGCGCTTTGATGCACTCTTTAATATCAA	780
QY	781	GGCAGTTGGTGGCGCTGCAAGTTTCTGGGTGTTGTTTCTATTGATGCTCCAGATAT	840
Db	781	GGCAGTTGGTGGCGCTGCAAGTTTCTGGGTGTTGTTTCTATTGATGCTCCAGATAT	840
QY	841	GGTCATGTTCTGGTGTGAGTGTGATCTTCTTCATCGCATTCGGCGCAGCGATTGC	900
Db	841	GGTCATGTTCTGGTGTGAGTGTGATCTTCTTCATCGCATTCGGCGCAGCGATTGC	900
QY	901	TTATGGCTCTTACTTGGTTCGCGCAACCGCAGCATTTGATCAGATGCAACCGCTCTCC	960
Db	901	TTATGGCTCTTACTTGGTTCGCGCAACCGCAGCATTTGATCAGATGCAACCGCTCTCC	960
QY	961	AGTGCCTGCAAGAACGACCAACCGCAAGCAGAACGCAACCGCAGAAATTTTCAACGATT	1020
Db	961	AGTGCCTGCAAGAACGACCAACCGCAAGCAGAACGCAACCGCAGAAATTTTCAACGATT	1020
QY	1021	CACCATCATCCAGGCACTTTGACCGGTGAAGCTATTGCACTGAGCAGCTGACGATGC	1080
Db	1021	CACCATCATCCAGGCACTTTGACCGGTGAAGCTATTGCACTGAGCAGCTGACGATGC	1080

QY	1081	CATGTTTGCAGCGGAAAGCTTGGCTCGGCGCTTGCCATCGTCCCAACCAAGGGCAGTT	1140
Db	1081	CATGTTTGCAGCGGAAAGCTTGGCTCGGCGCTTGCCATCGTCCCAACCAAGGGCAGTT	1140
QY	1141	AGTTTCTCCGGTGAAGTGGTGGCGATTCCCATCTGGCCATGCTTTTCGCAGT	1200
Db	1141	AGTTTCTCCGGTGAAGTGGTGGCGATTCCCATCTGGCCATGCTTTTCGCAGT	1200
QY	1201	TGCGACCAAGGCTGAGATGGTTCGAATGTGGATATCTTGATGCACATTTGGTTTCACAC	1260
Db	1201	TGCGACCAAGGCTGAGATGGTTCGAATGTGGATATCTTGATGCACATTTGGTTTCACAC	1260
QY	1261	AGTAAACCTCAACCGCACGCTTTAAACCGCTGGAAGAGCAGGGCGATGAAGTCAAAAGC	1320
Db	1261	AGTAAACCTCAACCGCACGCTTTAAACCGCTGGAAGAGCAGGGCGATGAAGTCAAAAGC	1320
QY	1321	AGGGAGCTGCTGTGTGAATTCGATATTTGATGCAATTAAGGCTGCAGGTTATGAGGTAAC	1380
Db	1321	AGGGAGCTGCTGTGTGAATTCGATATTTGATGCAATTAAGGCTGCAGGTTATGAGGTAAC	1380
QY	1381	CAGCCGATGTTGTTTTCGAATTTACAGAAACCGGACCTGTAAACACTTACGGTTTGG	1440
Db	1381	CAGCCGATGTTGTTTTCGAATTTACAGAAACCGGACCTGTAAACACTTACGGTTTGG	1440
QY	1441	CGAAATTTGAAGCGGAGCCAACTGCTCAACGTCGCAAGAAAGAAAGCGGTGCAGCAAC	1500
Db	1441	CGAAATTTGAAGCGGAGCCAACTGCTCAACGTCGCAAGAAAGAAAGCGGTGCAGCAAC	1500
QY	1501	ACCATAAGTTGAAACCTTGAGTTTCG	1527
Db	1501	ACCATAAGTTGAAACCTTGAGTTTCG	1527

RESULT 3
AP005282/c

LOCUS
DEFINITION
Corynebacterium glutamicum ATCC 13032 DNA, linear BCT 08-AUG-2002 9/10

ACCESSION
VERSION
AP005282 BA000036
AP005282.1 GI:21325287

KEYWORDS
SOURCE
ORGANISM

Corynebacterium glutamicum ATCC 13032
Corynebacterium glutamicum ATCC 13032
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Corynebacteriaceae; Corynebacterium.

REFERENCE

1
Nakagawa, S.
Complete genomic sequence of Corynebacterium glutamicum ATCC 13032
Unpublished
2 (bases 1 to 320550)
Nakagawa, S.
Direct Submission
Submitted (24-MAY-2002) Satoshi Nakagawa, Kyowa Hakko Kogyo Co.
Ltd., Tokyo Research Laboratories; 3-6-6, Asahi-machi, Machida,
Tokyo 194-8533, Japan (E-mail:snakagawa@xanagen.com,
Tel:81-44-829-3031, Fax:81-44-813-1651)
This sequence is conducted by collaboration of Kyowa Hakko Kogyo
Co. Ltd. And Kitasato University.

COMMENT

Co. Ltd. And Kitasato University.

FEATURES

source
1..320550
/organism="Corynebacterium glutamicum ATCC 13032"
/mol_type="genomic DNA"
/strain="ATCC 13032"
/db_xref="taxon:196627"
/note="ATCC 13032"
188..814
/gene="Cgl2516"
188..814
/gene="Cgl2516"
/codon_start=1
/transl_table=11
/product="Hypothetical protein"

```

/gene="BAB99909.1"
/db_xref="GI:21325288"
/translation="MYRAPDFVPLSMIVLVYTRKDRMETLAQAQTILLEKWCVAP
THASFVSIKAIPILSILLITLVVNGISSGNVPQPPALQOVRTDVVNKINIERKLK
GLVSIPELEHTAQTAQRNADSEKQPPDEGNLVIQONLPIYANADATIVDR
FLNSPDHVKLLANDYEAIGVAVYKGDHAWIVVEFTVAPADSVESBTE"
811..2151
/gene="Cgl12517"
811..2151
/gene="Cgl12517"
/codon_start=1
/transl_table=11
/product="Nicotinic acid phosphoribosyltransferase"
/protein_id="BAB99910.1"
/db_xref="GI:21325289"
/translation="NNTNPSEFSNRSTALLTDKELTWLQALADGSAERSTPEVF
SRRLPNERYGVAGTARVKAIRDFVTEBOLADLDFDDRTLEYLRYNRTQGVGD
YREGIYFQPSLLVGTGFAVCVILETVIISINADAVASAAAMVTAAADGREIIE
MGRSTHESYASVAGRAVLAGVFTTSNLEAARYGIPASGTSAAHWTLHINDDGTG
NEAAFKQAVRSLGVDITLLVDYDIITQGVATAIRVAGPDIGVRIISGDLGLARKV
RKQLDLNAHNTKIIVSSDLDEFAIAGRGEPVDFGVTSVTSGSAGTAGLVYKIV
EVAGHPVAKRSKSGYGGKKAIVTKRSGTAIEEIVYFPAEAPDTFGKLTLSLTI
PLMRDGEIVPEGLTLEDRAYLAKQLVSLPWEGLALSDEPVIHTRFVGFPPAA"
2438..2746
/gene="Cgl12518"
2438..2746
/gene="Cgl12518"
/codon_start=1
/transl_table=11
/product="Hypothetical protein"
/protein_id="BAB99911.1"
/db_xref="GI:21325290"
/translation="MGLRLNLPGRPOKTOITWKLPAQNRPOAPLSKOSNOTVNCFA
WLSSTPWLDAQCLRLNHPQIPQIRSMRPKSCLLPLKRWAVPVVLDKRWLR
"
2866..4608
/gene="Cgl12519"
2866..4608
/gene="Cgl12519"
/codon_start=1
/transl_table=11
/product="Rad3-related DNA helicases"
/protein_id="BAB99912.1"
/db_xref="GI:21325291"
/translation="MSTATIALQROLVNRPDLRLVDALEPLMERPTFAIMKGRSNYL
CMKVVARQELNQEDALIQEDIWLGKHIYRLNEWANETETGDRDDLDGVPYDLANK
QVSVARCEIGASCPHGEDCFATARGAKAEADVVTNHALDAIDALDSVLPEDHW
VVIDEAEHLDGRIITAVAGAEITVNSLNAARRAKLSDSKEERVQETAGDLETLQ
TMQGRWMDMSGKSTLVALKDALWALRAQIAGAEAGEAANDPERFAERQNLNHLM
ETHDANVRILEVFAEDPSKQVYVWHNDRRGSGLNVLPSVAGLHEKLEFAENTY
VLASATLTIGGNFNMAASWGLPKGSWDSMDAGTTPDPAKSGILYATRHLPDGRDL
PRETLEIYELLITAGGRTLGLFSSKRAARQATKAMRLLEPFDVLCQGDNTAALVKK
FSDSENTCLFGLTLWQGVDPGRSLSLVLDRIFFPPEDPDLLQARKRAADAEGRNG
FMEVAATHAALMAQAGELLRHVGDGRGVAVLDHRLSTKRYGGFLRFSNRPWETTN
PPTVRAALKRLVTK"
complement(4618..5511)
/gene="Cgl12520"
complement(4618..5511)
/gene="Cgl12520"
/codon_start=1
/transl_table=11
/product="Hypothetical membrane protein"
/protein_id="BAB99913.1"
/db_xref="GI:21325292"
/translation="MFSPIGSGQVGGDDNNNWSQGNPNSQBNQGYQAPYQOPYGQ
PFQGFNAYSSPIPEVPQPSQQAQRWRFDLGTVPQAKGFTATWQAWLSALIYF
AVLLWFMFAIPLMVSILVLAATSSGSDAAIAATGTSFFGFMIMVIAFISFVSLNC
YRNAARVGEQITIQSPFKMKGLKALGIYILINIVIFGIMLILIPGIIAAVLIF
AVPFAQLRDSIGDAFSAWKAWSKNVQGVILLELAIFALSFLGSAVILGMLVTPPL
TLLIYAYAFQTASGGPIMQOR"
complement(5587..6312)
/gene="Cgl12521"
complement(5587..6312)

```

```

/gene="Cgl12521"
/translation="MPKSLDLAHLKARASDGDSPSDPSTVQAMQIALHLPKQ
NPPRTDYLEAARSVKLCIDERVSTDPDFAALRWYGLHIRKVSRRANAANDRV
QDLPGVTVEDDAAKVRAFLPSAVVDVPADIKQLQISCTELPDLDENBPINDEPVYID
ESLKWTLGAAQAQVGHASMLLAHQPFWEVQNEAADFALHVRLEIPSEFLILIESPG
AVPRDGGFTVAPNTVTVAIP"
complement(6305..7324)
/gene="Cgl12522"
complement(6305..7324)
/gene="Cgl12522"
/EC_number="3.1.3.3"
/translation="PF00702:haloacid dehalogenase-like hydrolase
TIGR00099:TIGR00099: conserved hypothetical protein
TIGR00338:serB: phosphoserine phosphatase SerB"
/codon_start=1
/transl_table=11
/product="Phosphoserine phosphatase"
/protein_id="BAB99915.1"
/db_xref="GI:21325294"
/translation="MHGQSVVVELQETVQSSRPSRSHVVVVLGDPVDALDISRIGTTL
ADYDANIDTIRGISDPYPTGLELVKVTVPDVPSPGGEAMKLAALTSELNVDAIERS
GLLRSKLVCFDCDSTLITGEVIEMLAHAGKEAFAAATVTERAMRGELDPRESLRER
VKALAGDASVIDEVAALIELTPGARTTIRINNGYQTAUVVSGGFIQVLEGLAELE
LDYVRANTLEIVDGKLTGNVTGKIVDRAAKAEFLREFRADSLGKMWQTVAVGDGANDI
DMLSAAGLVAFNAKPAKLEIADTSVNHPLFDEVLHIMGISRDEIDLADQEDGTTTHRV
PLTNA"
complement(7811..9565)
/gene="Cgl12523"
complement(7811..9565)
/gene="Cgl12523"
/EC_number="1.9.3.1"
/translation="PF00115:Cytochrome c and Quinol oxidase polypeptide
1"
/codon_start=1
/transl_table=11
/product="Heme/copper-type cytochrome/quinol oxidases,
subunit 1"
/protein_id="BAB99916.1"
/db_xref="GI:21325295"
/translation="MTAVAPRVGDHVAQRPBPTGHARKSGKAWLMMTTTTHKQLGIM
YIIMSPSFYFGLGMLLIRALFLTPGLQFLSNEQNFQFLTMHGTVMLLLYGTPIVWG
FANYVLPQIGADPVAFPRLNAGFWITTVGVAMITGLTTPGGAADFQWTFPLSD
AIHSPGLGSDMWIVGVGATIGSVASAINMLTTLICLRAPGTMFRMPITWNIPVVS
VIALLIIFLLAALGLVLYDRKLGHLDPANGSLLOHLEFWFQGHPPVYVIALPFF
GIVSIIIPVSRKPMFVGLLIFATLISGALSNVAMHMFVTGAVLLPFPFPMFLFI
SVPTGVKPFNVGTWKGHITWETPMINSVGMATFLGGLTGIMLASPLDPLHLADS
YFLLAHFYTTLPGTVFASCAGVYFPMGTGEMMDERLGTIHWLTFVYFGHTGLIQ
HWNGMMPRRYADYLDSDGFTLYNQISTVFSFLGLSLVPIFINVWFKSWRRTGELTV
DDPWYGNLSLEWATSCPPPHNPFASLPIRISRPAPFLHYPHMIERMAEAHTGHHDH
INAPBLGTAPALADSSR"
9701..10009
/gene="Cgl12524"
9701..10009
/gene="Cgl12524"
/codon_start=1
/transl_table=11
/product="Hypothetical protein"
/protein_id="BAB99917.1"
/db_xref="GI:21325296"
/translation="MPKKNVFTIHEVAARFSLHNAVTFNQRCEGVHNLTDISPGSPT
PTVNPCHRFHNSPRGSIWLRPFQGRFLLRKDPVPPRPGRAWRSPTPICASEKPAF
"
complement(10086..11090)
/gene="Cgl12525"
complement(10086..11090)
/gene="Cgl12525"
/translation="PF00268:Ribonucleotide reductase, small chain"

```


		/codon_start=1	
		/transl_table=11	
		/product="Ribonucleotide reductase beta subunit"	
		/protein_id="BAB99918.1"	
		/db_xref="GI:21325297"	
		/translation="MAADSDLSVHDAYLKEHVAPVKAINWNSIPDSKDLVMDRLGN FWLPEKVPVNSDIKSGTNEVEKAAATMRVFTGLTLLDTIOGTGVAISLLPADSLHE EALVNIAPMESVHAKSYNSIEMTLASTAEINDAFRWSBENENLORAKIILSYBGD DPLRKIASVILLESFLFYSGFYLPMWSSHSKLTADVTRLIRDEAVHGYVIGKY QKAVAKETPEROBELEKTFDLYLDYDNETOYSEDLVLDLGTWEDVKERFLRYNANKA LNNLVEGELFPADETKVSNNILSALSFNADENHDFSGSGSSYVIKAENTEDDDWDF "	
gene		11555..12043	
		/gene="Cgl2526"	
CDS		11555..12043	
		/gene="Cgl2526"	
		/note="PF00210: Ferritin	
Query Match		99.2%; Score 1515; DB 1; Length 320550;	
Best Local Similarity		99.9%; Pred. No. 0;	
Matches 1526; Conservative		0; Mismatches 0; Indels 1; Gaps 1;	
QY	1	CTCATCGCATCTCGCGCCTTCGGTCTTCGCCAGTGTGGTTGCTTACCGGCAACCAAG	60
Db	148077	CTCATGCGCATCTCGCGCGTTCGGGTTCTTGCCAGTGTGGTTGCTTACCGGCAACCAAG	148018
QY	61	CGTTTCGGCGGCAATGAGTTCTTGGCGCGCGGTATTGGTATGGCGATGGTGTCCCGAG	120
Db	148017	CGTTTCGGCGGCAATGAGTTCTTGGCGCGCG-GTATTGGTATGGCGATGGTGTCCCGAG	147959
QY	121	CTTGTGTGAACGGCTACGACGTGGCGCGCACCATGCTGGCGGCGAAATGCCAATGTGTC	180
Db	147958	CTTGTGTGAACGGCTACGACGTGGCGCGCACCATGCTGGCGGCGAAATGCCAATGTGTC	147859
QY	181	CGTGTTCGTTTGTAGATGTGCCAAGCCGTTTACGAGGCGACGTCGTCCTGTGCTGGT	240
Db	147898	CGTGTTCGTTTGTAGATGTGCCAAGCCGTTTACGAGGCGACGTCGTCCTGTGCTGGT	147859
QY	241	GGTTTCTTGGAATCTCGGCAACGATCGAGAAAGTTCTCTGCAACGCACTCAAGGGCACTGC	300
Db	147838	GGTTTCTTGGAATCTCGGCAACGATCGAGAAAGTTCTCTGCAACGCACTCAAGGGCACTGC	147779
QY	301	AGACTTCCTGATCACTCCAGTGTGACGTTGCTGCTCACCGGATTCCTTACATTCATCGC	360
Db	147778	AGACTTCCTGATCACTCCAGTGTGACGTTGCTGCTCACCGGATTCCTTACATTCATCGC	147719
QY	361	CATTGCGCCAGCAATCGCTGGTGGCGGATGTCTGGCACAGGCTACAGGGACTTTTA	420
Db	147718	CATTGCGCCAGCAATCGCTGGTGGCGGATGTCTGGCACAGGCTACAGGGACTTTTA	147659
QY	421	TGATTTCCGTTGTCAGTCCGGCGGTCTGCTCTTCGGTCTGGTCTACTCAACCAATCGTCAT	480
Db	147658	TGATTTCCGTTGTCAGTCCGGCGGTCTGCTCTTCGGTCTGGTCTACTCAACCAATCGTCAT	147599
QY	481	CACGTGTCGACACGTCCTTCCCGCAATGAGCTGAGCTGAGCTGTTTAAACAGGGTGATC	540
Db	147598	CACGTGTCGACACGTCCTTCCCGCAATGAGCTGAGCTGAGCTGTTTAAACAGGGTGATC	147539
QY	541	CTTCATCTTCGCAACCGGATCTATGCTATATATCGCCACAGGTCGCGCATGTTTGGCAGT	600
Db	147538	CTTCATCTTCGCAACCGGATCTATGCTATATATCGCCACAGGTCGCGCATGTTTGGCAGT	147479
QY	601	GTTCCTTCCTGGCGAAGAGTGAAGAGTCAAGGCGCTTGCAGGTCCTTCAGGTCTCCCG	660
Db	147478	GTTCCTTCCTGGCGAAGAGTGAAGAGTCAAGGCGCTTGCAGGTCCTTCAGGTCTCCCG	147419
QY	661	TGTTCTTGTTATPACGGAGCTCCGATCTTTCGGTGTGAACCTTCGCGTCCGCTGGCGGTT	720
Db	147418	TGTTCTTGTTATPACGGAGCTCCGATCTTTCGGTGTGAACCTTCGCGTCCGCTGGCGGTT	147359
QY	721	CTTCATCGGTATCGGTACCGCAGCTATCGGTGGCGCTTGTGATTCGACTCTTTAATATCAA	780
Db	147358	CTTCATCGGTATCGGTACCGCAGCTATCGGTGGCGCTTGTGATTCGACTCTTTAATATCAA	147299

QY	781	GGCAGTTGCGTTGGCGCGCTGCAGGTTTCTTGGGTGTTGTTTCTATTGATGCTCCAGATAT	840
Db	147298	GGCAGTTGCGTTGGCGCGCTGCAGGTTTCTTGGGTGTTGTTTCTATTGATGCTCCAGATAT	147239
QY	841	GGTCATGTTCTTGGTGTGTGTCAGTTGTTACCTTCTTTCATCGCATTCGGCGCAGGATTGC	900
Db	147238	GGTCATGTTCTTGGTGTGTGTCAGTTGTTACCTTCTTTCATCGCATTCGGCGCAGGATTGC	147179
QY	901	TATGCGCTTACTTGGTTCGCGCGAAACCGCAGCATTTGATCCAGATGCAACCGCTGCTCC	960
Db	147178	TATGCGCTTACTTGGTTCGCGCGAAACCGCAGCATTTGATCCAGATGCAACCGCTGCTCC	147119
QY	961	AGTGCTTCGAGGAAACACCAAGCCGAGCAGAAAGCACCGCAGAAATTTTCAACCAATTC	1020
Db	147118	AGTGCTTCGAGGAAACACCAAGCCGAGCAGAAAGCACCGCAGAAATTTTCAACCAATTC	147059
QY	1021	CACCATCATCCAGGCACCTTTTTCACCGGTGAAGCTATTGCACTGAGCAGGTTCAGCGATGC	1080
Db	147058	CACCATCATCCAGGCACCTTTTTCACCGGTGAAGCTATTGCACTGAGCAGGTTCAGCGATGC	146999
QY	1081	CATGTTTGCAGCGGAAAGCTTGGCTCGGGCGTGGCCATCGTCCCAACCAAGGGGCAATT	1140
Db	146998	CATGTTTGCAGCGGAAAGCTTGGCTCGGGCGTGGCCATCGTCCCAACCAAGGGGCAATT	146939
QY	1141	AGTTTCTCCGGTGAGTGGAAAGATTGCTGTGGCATTCCTCATCTGGCCATGCTTTCCAGT	1200
Db	146938	AGTTTCTCCGGTGAGTGGAAAGATTGCTGTGGCATTCCTCATCTGGCCATGCTTTCCAGT	146879
QY	1201	TGCGACCAAGGCTGAGGATGGTTCCAAATGTGGATATCTTGAATGACATTTGGTTTCGACAC	1260
Db	146878	TGCGACCAAGGCTGAGGATGGTTCCAAATGTGGATATCTTGAATGACATTTGGTTTCGACAC	146819
QY	1261	AGTAACTCTCAACGGCACCGACTTTTAAACCGCTGGAAGAGCAGGCGATGAAGTCAAAGC	1320
Db	146818	AGTAACTCTCAACGGCACCGACTTTTAAACCGCTGGAAGAGCAGGCGATGAAGTCAAAGC	146759
QY	1321	AGGGGAGCTGCTGTGTGAATTCGATATTCGATGCAATTAAGGCTGCAGGTTATCAGGTTAAC	1380
Db	146758	AGGGGAGCTGCTGTGTGAATTCGATATTCGATGCAATTAAGGCTGCAGGTTATCAGGTTAAC	146699
QY	1381	CAGCGCGATTTGTTTTCGAAATTAAGAAACCGGACCTGTAAACACATTTACCGTTTGGG	1440
Db	146698	CAGCGCGATTTGTTTTCGAAATTAAGAAACCGGACCTGTAAACACATTTACCGTTTGGG	146639
QY	1441	CGAAATTAAGCGGGAGCGCAACCTGCTCAACGCTGCGCAAGAAAGAAAGCGTGCAGCAAC	1500
Db	146638	CGAAATTAAGCGGGAGCGCAACCTGCTCAACGCTGCGCAAGAAAGAAAGCGTGCAGCAAC	146579
QY	1501	ACCATAGTTTGAACCTTCAGTGTTCG 1527	
Db	146578	ACCATAGTTTGAACCTTCAGTGTTCG 146552	

RESULT 4	
LOCUS BX927155/c	349136 bp DNA linear BCT 10-JUN-2004
DEFINITION Corynebacterium glutamicum ATCC 13032, IS fingerprint type 4-5, complete genome; segment 8/10.	
ACCESSION BX927155	
VERSION BX927155.1	GI:41326514
KEYWORDS complete genome.	
SOURCE Corynebacterium glutamicum ATCC 13032	
ORGANISM Corynebacterium glutamicum ATCC 13032	
REFERENCE 1 (bases 1 to 349136)	
AUTHORS Kalinowski, J., Bathe, B., Bartels, D., Bischoff, N., Bott, M., Burkovski, A., Dusch, N., Eggeling, L., Eikmanns, B. J., Gaigalat, I., Goessmann, A., Hartmann, M., Hutmacher, K., Kramer, R., Linke, B., McHardy, A. C., Meyer, F., Mockel, B., Pfeifferle, W., Puhler, A., Rey, D. A., Ruckert, C., Rupp, O., Sahn, H., Wendisch, V. F., Wiegand, I. and Tauch, A.	

TITLE The complete Corynebacterium glutamicum ATCC 13032 genome sequence and its impact on the production of L-aspartate-derived amino acids and vitamins

JOURNAL J. Biotechnol. 104 (1-3), 5-25 (2003)

MEDLINE 22830012

PUBMED 12948626

REFERENCE 2 (bases 1 to 349136)

AUTHORS Kalinowski, J

TITLE Direct Submission

JOURNAL Submitted (21-JAN-2004) Joern Kalinowski, Institut fuer Genomforschung, Universitaet Bielefeld; Universitaetstrasse 25, 33615 Bielefeld, Germany

COMMENT E-mail:Joern.Kalinowski@ceitec.uni-bielefeld.de

This sequence was accomplished by collaboration between Degussa AG and Bielefeld University.

Join(BX927148.1:1..348071,BX927149.1:51..349887, BX927150.1:51..348475, BX927151.1:51..349459, BX927152.1:51..349799, BX927153.1:51..349584, BX927154.1:51..349575, BX927155.1:51..349136, BX927156.1:51..349115, BX927157.1:51..140057)

FEATURES

Location/Qualifiers

1..349136

/organism="Corynebacterium glutamicum ATCC 13032"

/mol_type="genomic DNA"

/strain="DSM 20300 = ATCC 13032"

/db_xref="taxon:196627"

/note="IS fingerprint type: 4-5"

complement(67..729)

/locus_tag="cg2562"

complement(67..729)

/locus_tag="cg2562"

/codon_start=1

/transl_table=11

/product="Serine/threonine specific protein phosphatase"

/protein_id="CAF20997.1"

/db_xref="GI:41326515"

/translation="WYKQLDSLVFTDGESIAKARLASMTDMWFSDHLHGKHFVASM RGFDDPDRHDEVILSNFNTIGDDVLTGLDSSGAHRAERLGLAERLGGVVKH LVPGNDSCHPMRYHAYKREPLEVFDVSOAFQRMKWDDEVDVLSHPRPQDHPGM ESRFDDRLPFLIHGHLHSQFPMTGQDVGVGAEWGLKFAPRELVQLKLWESLSE Kt"

830..2722

/gene="lccop"

/locus_tag="cg2563"

830..2722

/gene="lccop"

/locus_tag="cg2563"

/codon_start=1

/transl_table=11

/product="ECTOINE BETAINE TRANSPORTER"

/protein_id="CAF20998.1"

/db_xref="GI:41326516"

/translation="WSTNSGNLPSQSPSEPHPHDTHPGLVPGISVDAQRNKFGL DKTVFGTAALIFATFIANGISDPSSVSTSMFSWAMTGMLLNFVMLIGIGTMY IAFSRYKLGTDDEPFESFWIAMFGAGIVGVPFPGSEPLHYLSPPPHYT EGSTPSLHQALQGHFHWGLSAGLVAGALAYSRYRRGRVTLISSTFRSLFGKK TEGIAGRLDMWAIATLPGTAATGLSAIQVCGQVQIISGASEITNNILIAIILIT TGFIISSVCSVKGRILSNLINSITLGLVLVFTVGTPIFLNLIIPSVLYGSEFL SWAGLSWGEETIFQAGATPYAWMTIAWTFVGMFTIARISRGTLREFAITWAI PSFIIILATPIFGGTATNRRNVDPGSSKEQVLFDMFNSLPLSYTPIFILFVL AVFFVTSDASVSMGTSSQGNPAPNKLIVVFWGLGMMGIAMVMLLTGSESLTGL NUTIIAIPFALVIAIWMATFIKDLSTDPAAIRQRYAKAAISNAVVRVGLIEHGDDEL STEPABEGRGAGATFDSADHTIDWYQRTDEBGNVDVYDFTTGKADGWTPSTEEGE VDAKD"

complement(3014..3295)

/locus_tag="cg2564"

complement(3014..3295)

/locus_tag="cg2564"

/codon_start=1

/transl_table=11

/product="hypothetical protein predicted by Glimmer/Critica"

/protein_id="CAF20999.1"

db_xref="GI:41326517"

/translation="WATTASKISTIRPAQDALMSVREDLHARFDGLVDPVQVDAILD HVASNRRAKITVFSKIPIAREATAALQIAGNVNADLLDFIALNRGMAA"

3674..4672

/locus_tag="cg2565"

3674..4672

/locus_tag="cg2565"

/codon_start=1

/transl_table=11

/product="hypothetical protein predicted by Glimmer/Critica"

/protein_id="CAF21000.1"

/db_xref="GI:41326518"

/translation="WNPETTTIDRISKEQDPAARSRVQFIVETVRALPNLTTKQAS LAIOLLDAVQLADAAGTKGASTSNASSLPDTFDALTSILGKLDVRSDESWRSFGFQF SETAHPMLIAPEIEIFQHTDVRPGSDDAVAPQENQDMWRRLGSVTRPNLYKE FSGQKQARAVEMLGNLWKIIVVSYNTESFRLKGLTRVRYTTPGHPVPLMKCKWY SIRVSTIGENQPEIVRCLGEIFCGYLFQMWLKEPVRKAGLRKIQESAAAYIAWARL DLSPTGNTTNTNSYISTRPLSPAPRWDVLEASHQLENLRLRGDTGPTVATQSAAG"

4836..6431

/locus_tag="cg2566"

4836..6431

/locus_tag="cg2566"

/codon_start=1

/transl_table=11

/product="putative secreted protein"

/protein_id="CAF21001.1"

/db_xref="GI:41326519"

/translation="MKRNLLVAPLTASLVFCNLAVANAWEVREASPVVINEVSNLS DPGDWBELANTNNNSIDISGSLVDDKEDLENALVPEGETEISGGGFVIYTDGAD YVPTNTPFGQGEYFGLGKDDTVLTENAEVVATYSWKDLGHEAENTYGRIPDMTGDFA ANTGVPTGAKNVAEGSGEBEGVANAQLPFHNVEITPIHLGGDTGEDMSGVDFDA NSTAWIANDIKTIVSLAHDIANNITKLTGEWETGYEGGGEPEADEGTVAANGDIYL STERNANDKVSRSILSFATPTGKTGVQNAQVQWDLSEFVGDIOPNGSLRAIAOLED NIFVGVETGDIVVDLSAQDQVLVQRYESSFDGWSLDYNAATKQLSVUCDEACDG LSEILWEPGRKLYKSDDKIERPANLGNWAEFGTYISELCKENGVSVSYSLWAD DAATNEGTSLSAQVINGDCGVNIPEGSSDNSSDFATGSIAGATAVAVLVGIIA GALGGFFQIILAAFPALQQVIRF"

6556..6942

/locus_tag="cg2567"

6556..6942

/locus_tag="cg2567"

/codon_start=1

/transl_table=11

/product="hypothetical protein predicted by Glimmer/Critica"

/protein_id="CAF21002.1"

/db_xref="GI:41326520"

/translation="NPPSFFPRRRKSEPDLLALPAQIRDVHADVLIDIFMAESLVIINI DEKSAKLIDAAARHHIPRTFTGPNARPLSVIPIEDPRSRPTLHPDHGWMPLSPFVVD ELLGGGLKIGTELESTNIAFTVDAS"

complement(7072..8385)

/gene="dctM"

/locus_tag="cg2568"

complement(7072..8385)

/gene="dctM"

/locus_tag="cg2568"

/codon_start=1

/transl_table=11

/product="PUTATIVE C4-DICARBOXYLATE TRANSPORT SYSTEM (PERMEASE LARGE PROTEIN)"

/protein_id="CAF21003.1"

/db_xref="GI:41326521"

/translation="MLSPAVAALILVIGIVVLIASVFAVAIGLPSLFAAMAVLGP ENAAQVQRMFTGNTSPTLLAIPEFVLGILMNSGGIATRLIDAQVLVGRMPASMA NNTNIANGLFAGVSGAAVASAGVTMTKPKRGYSRAYAAVNAVVASAPAGMLIPP SNTFIYLSVSTSAALFMAGVGGGLLWILACVIVGTWLARKENYKREIHTPTFKQS LVLVWRLPSSLMIIVVGGILLGMWFTPTESAIAIVVYLCVIGFYRTIKTVGLDAIL LKATRTTIVMLLIAVSALWSMAFAKIPQMSIDALLSVDSKIVLLIMFILLILI GTVMDPTPAIIFVPIFVTVTELGVDPVHFQAMVNMVLSGVITPVGNVLFVGSQV AGLRVETVIRRLWPYLIATIALVFFVVFVFPQISILWLPPTMGLMG"

complement(8387..8911)

/gene="dctX"

```
CDS
/locus_tag="cg2569"
/complement(8387..8911)
/gene="dctx"
/locus_tag="cg2569"
/codon_start=1
/transl_table=11
/product="PUTATIVE C4-DICARBOXYLATE TRANSPORT SYSTEM
(PERMEASE SMALL PROTEIN)"
/protein_id="CAF21004.1"
/db_xref="GI:41326522"
/translation="MNNEFKSIVCVTAQWVFSROVLHSPSTWSELSKLLFWMLSPAGS
ALFLGKRGHIAVDFIARKLPVSAQRVLQVLIIVFALLGIMWGGYLAASIAMNQ
LFLALPLTLGWYVIVPIAFIADLIIEVATGKERPYPLVDSERPRDLDELA
QSAIDSASAEGRN"
/complement(8908..9657)
/gene="dctp"
/locus_tag="cg2570"
/complement(8908..9657)
/gene="dctp"
/locus_tag="cg2570"
/codon_start=1
/transl_table=11
/product="PUTATIVE C4-DICARBOXYLATE-BINDING PROTEIN"
/protein_id="CAF21005.1"
/db_xref="GI:41326523"
/translation="MMGSVEKKYLRFDWFTTPEVGVIGSIHQMVSVIRHODLMEPL
FOSLADNISVIGFTQGRNLYTDDAPVKRPADLAGKIRVOESAMHRLMELMG
SATPLTYGVVYTAQMSGLVDKENNEISVYTONHEFVARVNSNTHLVGLDYVWRHD
LLDAMSEPRERELFLEWDADMEHTDLNWTETDAVIEKAKAGAEFVEVDAQFTDAL
APIKDFLTSFEQREYIEAVRAADTSGAAS"
/complement(9722..11569)
/gene="lepA"
/locus_tag="cg2571"
/complement(9722..11569)
/gene="lepA"
/locus_tag="cg2571"
/codon_start=1

Query Match          99.2%; Score 1515; DB 1; Length 349136;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1526; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 CTCATGCGAFCGCGCGTTCGCTTTCGCCAGTGTGGTTCACCGCAACCAAG 60
DB 340135 CTCATGCGCATCTGCGCGTTCGCGTTCGCCAGTGTGGTTCACCGCAACCAAG 340076

QY 61 CGTTTCGGGGCAATCAGTTCCTGGGGCGCGCTATTGGTATGCGCATGGTTCGCGAG 120
DB 340075 CGTTTCGGGGCAATCAGTTCCTGGGGCGCGG-GTATTGGTATGCGCATGGTTCGCGAG 340017

QY 121 CTTGGTGAACGGCTACGACGTGGCGCGCCACCATGGCTGGGGCGAAATGCCAATGTGGTC 180
DB 340016 CTTGGTGAACGGCTACGACGTGGCGCGCCACCATGGCTGGGGCGAAATGCCAATGTGGTC 339957

QY 181 CCGTTTGGTTTAGATTGTTGCCAACCGGTTTACGAGGCAACCGTCTTCCTGCTGGT 240
DB 339956 CCGTGTGTGGTTTAGATTGTTGCCAACCGGTTTACGAGGCAACCGTCTTCCTGCTGGT 339897

QY 241 GGTTCCTTGGATTCTGGCAACGATCGAGAGTTCTTCGCAACGCACTCAAGGGCACTGC 300
DB 339896 GGTTCCTTGGATTCTGGCAACGATCGAGAGTTCTTCGCAACGCACTCAAGGGCACTGC 339837

QY 301 AGACTTCCTGATCAGTCCAGTGTGAGTGTGCTGCACCGGATTCCTTACATTCATCGC 360
DB 339836 AGACTTCCTGATCAGTCCAGTGTGAGTGTGCTGCACCGGATTCCTTACATTCATCGC 339777

QY 361 CATTTGCCCGCAGCAATCGCTGGGTGGGCGATGTGCTGGCACACGCTACAGGGAATTTA 420
DB 339776 CATTTGCCCGCAGCAATCGCTGGGTGGGCGATGTGCTGGCACACGCTACAGGGAATTTA 339717

QY 421 TGAATTCGGTGTCCAGTCGGCGGCTCTGCTCTTCGGTCTGGTCTACTACCAATGCTCAT 480
DB 339716 TGAATTCGGTGTCCAGTCGGCGGCTCTGCTCTTCGGTCTGGTCTACTACCAATGCTCAT 339657
```

```
QY 481 CACTGTGCTGCACCAAGTCCTTCCCGCCAAATTGAGCTGGAGCTCTTTAAACCAAGGTGGATC 540
DB 339656 CACTGTGCTGCACCAAGTCCTTCCCGCCAAATTGAGCTGGAGCTCTTTAAACCAAGGTGGATC 339597

QY 541 CTTTCATCTTTGCGCAACGGCATCTATAGCTAATATCGCCACAGGGTGCAGCATGTTTGGCAGT 600
DB 339596 CTTTCATCTTTGCGCAACGGCATCTATAGCTAATATCGCCACAGGGTGCAGCATGTTTGGCAGT 339537

QY 601 GTTCTTCTCTGGCGAAGAGTGAAGAAGCTCAAGGGCTTTCAGAGTGCCTTCAGGTGTCTCCGC 660
DB 339536 GTTCTTCTCTGGCGAAGAGTGAAGAAGCTCAAGGGCTTTCAGAGTGCCTTCAGGTGTCTCCGC 339477

QY 661 TGTTCCTTGGTATTATACCGAGCTCTCGATCTTTCGGTGTGAACCTTCGCTCGCTGGCGCGTT 720
DB 339476 TGTTCCTTGGTATTATACCGAGCTCTCGATCTTTCGGTGTGAACCTTCGCTCGCTGGCGCGTT 339417

QY 721 CTTTCATCGGTATCGGTACCGCATCTATCGGTGGCGCTTTGATTCGATCTCTTTAATATCAA 780
DB 339416 CTTTCATCGGTATCGGTACCGCATCTATCGGTGGCGCTTTGATTCGATCTCTTTAATATCAA 339357

QY 781 GGCAGTTGCGTTGGGCGCTGCAGGTTTCTTGGTGTGTTTCTTATTGATCTCCAGATAT 840
DB 339356 GGCAGTTGCGTTGGGCGCTGCAGGTTTCTTGGTGTGTTTCTTATTGATCTCCAGATAT 339297

QY 841 GGTTCATGTTCTTGTGTGTGTCAGTGTGTTACCTTCTTCATCGCATTCGGCGCAGGATTGC 900
DB 339296 GGTTCATGTTCTTGTGTGTGTCAGTGTGTTACCTTCTTCATCGCATTCGGCGCAGGATTGC 339237

QY 901 TTATGGCCCTTTACTTGGTTTCGGCGCAACGGCAGCATTTGATCCAGATGCCAACCCCTGCTCC 960
DB 339236 TTATGGCCCTTTACTTGGTTTCGGCGCAACGGCAGCATTTGATCCAGATGCCAACCCCTGCTCC 339177

QY 961 AGTGCCCTGCAGGAACGACCAAGCCGAGCAAGACACCCGCGAGATTTTCAAACGATTC 1020
DB 339176 AGTGCCCTGCAGGAACGACCAAGCCGAGCAAGACACCCGCGAGATTTTCAAACGATTC 339117

QY 1021 CACCATCATCCAGGCACCTTTTGACCGGTGAAGCTATTGGACTGACAGCGTCAAGCGATGC 1080
DB 339116 CACCATCATCCAGGCACCTTTTGACCGGTGAAGCTATTGGACTGACAGCGTCAAGCGATGC 339057

QY 1081 CATGTTTGGCAGCGGAAAGCTTGGCTCGGGCGTGGCCATGTCCTCCAAACCAAGGGGCGAGTT 1140
DB 339056 CATGTTTGGCAGCGGAAAGCTTGGCTCGGGCGTGGCCATGTCCTCCAAACCAAGGGGCGAGTT 338997

QY 1141 AGTTTCTCCGGTGAAGTGAAGAAGATTGTTGGTGGCATTTCCCATCTGGCCATGCTTTCCAGT 1200
DB 338996 AGTTTCTCCGGTGAAGTGAAGAAGATTGTTGGTGGCATTTCCCATCTGGCCATGCTTTCCAGT 338937

QY 1201 TCGCACCAAGGCTGAGGATGGTTCCCAATGTGGATATCTTTGATGCACATTGGTTTCCGACAC 1260
DB 338936 TCGCACCAAGGCTGAGGATGGTTCCCAATGTGGATATCTTTGATGCACATTGGTTTCCGACAC 338877

QY 1261 AGTAAACCTTCAACGGCACCGCATTTTAAACCCGCTGGAAGAGAGAGCGGCGATGAAGTCAAAGC 1320
DB 338876 AGTAAACCTTCAACGGCACCGCATTTTAAACCCGCTGGAAGAGAGAGCGGCGATGAAGTCAAAGC 338817

QY 1321 AGGGGAGCTGCTGTGTGAATTCGATATTGATGCGCATTAAGGCTGCAGGTTATGAGGTAAC 1380
DB 338816 AGGGGAGCTGCTGTGTGAATTCGATATTGATGCGCATTAAGGCTGCAGGTTATGAGGTAAC 338757

QY 1381 CAGCGCGATTGTTGTTTTCGAATTTACAAGAAAACCGGACCTGTAAACACACTTACGTTTGGG 1440
DB 338756 CAGCGCGATTGTTGTTTTCGAATTTACAAGAAAACCGGACCTGTAAACACACTTACGTTTGGG 338697

QY 1441 CGAAATTTGAAGCGGAGCGCAACCTGCTCAACCGTTCGCAAAAGAAAGACCGGTGCCAGCAAC 1500
DB 338696 CGAAATTTGAAGCGGAGCGCAACCTGCTCAACCGTTCGCAAAAGAAAGACCGGTGCCAGCAAC 338637

QY 1501 ACCATAAGTTGAACCTTGAGTGTTCG 1527
DB 338636 ACCATAAGTTGAACCTTGAGTGTTCG 338610
```


PI OZAKI
PC C12N15/09, C12N15/09, C07K14/34, C07K16/12, C07K16/40, C12M1/00, PC
C12N1/15,
PC C12N1/19, C12N1/21, C12N5/10, C12N9/00, C12N9/02, C12P7/40, C12P13/
04, C12P13/08,
PC C12P19/00, C12P19/34, C12P21/02, C12Q1/37, C12Q1/68, G01N33/53, PC
G01N33/566,
PC G01N33/569, G01N33/58, G01N37/00, C12P21/08, C12N1/21, C12R1/15,
C12N1/21, C12R1/13, C12N1/21, C12R1/01, C12P13/08, C12R1/15,
PC C12N15/00,
PC C12N5/00, C12N15/00
CC Novel polynucleotide
FH Key Location/Qualifiers
FT source 1..1983
FT source /organism='Corynebacterium glutamicum'.
FEATURES
source
1..1983
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
ORIGIN
Query Match 97.7%; Score 1492; DB 6; Length 1983;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1503; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1 CTCATGGCATCTCGCGCGTTCGGTCTTCCAGTGTGGTGGTTTCACCGCAACCAAG 60
DB 481 CTCATGGCATCTCGCGCGTTCGGTCTTCCAGTGTGGTGGTTTCACCGCAACCAAG 540
QY 61 CGTTTCGGCGCAATGAGTTCCTGGCGCGCGTATTTGATGCGCATGGTGTCCCGAG 120
DB 541 CGTTTCGGCGCAATGAGTTCCTGGCGCGG-CTATTGATGCGCATGGTGTCCCGAG 599
QY 121 CTTGTTGAACGGCTACGACGTGCGCCACCATGTGCTCGCGGCGAAATGCCAATGTGTC 180
DB 600 CTTGTTGAACGGCTACGACGTGCGCCACCATGTGCTCGCGGCGAAATGCCAATGTGTC 659
QY 181 CTTGTTGTTTGTAGTGTGCGCAACCGGTTTACAGGCGACCGTCTTCTGTGCTGT 240
DB 660 CTTGTTGTTTGTAGTGTGCGCAACCGGTTTACAGGCGACCGTCTTCTGTGCTGT 719
QY 241 GGTTCCTTGGATTCTGGCAACGATCGAAGAGTTCTTCGCAACGCACTCAAGGGCACTGC 300
DB 720 GGTTCCTTGGATTCTGGCAACGATCGAAGAGTTCTTCGCAACGCACTCAAGGGCACTGC 779
QY 301 AGACTTCCTGATCACTCCAGTGTGCTGCTTCTGCTCAGCGGATTCCTTACATTCATCGC 360
DB 780 AGACTTCCTGATCACTCCAGTGTGCTGCTTCTGCTCAGCGGATTCCTTACATTCATCGC 839
QY 361 CATTTGCCCCAGCAATCGCTGGTGGCGATGTCGACACGCGTCTACAGGGACTTTA 420
DB 840 CATTTGCCCCAGCAATCGCTGGTGGCGATGTCGACACGCGTCTACAGGGACTTTA 899
QY 421 TGATTTCCGTTGCTCCAGTCCGCGGTCTGCTCTTCCGTTCTGCTTCTGCTTCTGCTCAT 480
DB 900 TGATTTCCGTTGCTCCAGTCCGCGGTCTGCTCTTCCGTTCTGCTTCTGCTTCTGCTCAT 959
QY 481 CACTGTCTGCAACGATCTTCCCGCAATTGAGCTGGAGCTGTTTAAACAGGGTGGATC 540
DB 960 CACTGTCTGCAACGATCTTCCCGCAATTGAGCTGGAGCTGTTTAAACAGGGTGGATC 1019
QY 541 CTTTCATCTTCGCAACGCGATCTATGCTAATATCGCCACGAGGTGGCGCATGTTGGCAGT 600
DB 1020 CTTTCATCTTCGCAACGCGATCTATGCTAATATCGCCACGAGGTGGCGCATGTTGGCAGT 1079
QY 601 GTTCTTCTGGCGAAGAGTGAAGCTCAAGGCCCTTGCAGGTGCTTCAAGGTCTCCGC 660
DB 1080 GTTCTTCTGGCGAAGAGTGAAGCTCAAGGCCCTTGCAGGTGCTTCAAGGTCTCCGC 1139
QY 661 TGTCTTGTGTTATACGAGGCTCGGATCTTCGTTGTAACCTTCGCTGGCTGGCGGTT 720
DB 1140 TGTCTTGTGTTATACGAGGCTCGGATCTTCGTTGTAACCTTCGCTGGCTGGCGGTT 1199

QY 721 CTTTCATCGGTATCGGTACCGAGCTATCCGTTGGCGCTTTGATGTCACCTCTTTAATATCAA 780
DB 1200 CTTTCATCGGTATCGGTACCGAGCTATCCGTTGGCGCTTTGATGTCACCTCTTTAATATCAA 1259
QY 781 GGCAGTTCGTTGGCGCTGTCAGGTTCTTGGGTGTTGTTCTTATGATGTCGCAAGATAT 840
DB 1260 GGCAGTTCGTTGGCGCTGTCAGGTTCTTGGGTGTTGTTCTTATGATGTCGCAAGATAT 1319
QY 841 GGTCAATGTTCTTGGTGTGTCAGTGTGTACCTTCTTCATCGCATTCGCGGCGAGCGATTGC 900
DB 1320 GGTCAATGTTCTTGGTGTGTCAGTGTGTACCTTCTTCATCGCATTCGCGGCGAGCGATTGC 1379
QY 901 TTATGGCCTTTACTTGGTTTCGGCCACCGCAGCATTTGATCCAGATGCAACCGCTGCTCC 960
DB 1380 TTATGGCCTTTACTTGGTTTCGGCCACCGCAGCATTTGATCCAGATGCAACCGCTGCTCC 1439
QY 961 AGTCTGTCGAGAACGACCAAGCCGAGCAGACACCGCAGAAATTTTCAACGATTC 1020
DB 1440 AGTCTGTCGAGAACGACCAAGCCGAGCAGACACCGCAGAAATTTTCAACGATTC 1499
QY 1021 CACCATCATCAGGCAACCTTTGACCGGTGAAGCTATTGCACTGAGCAGCGTCAGCGATGC 1080
DB 1500 CACCATCATCAGGCAACCTTTGACCGGTGAAGCTATTGCACTGAGCAGCGTCAGCGATGC 1559
QY 1081 CATGTTTCCAGCGCAAAAGCTTGGCTCGGGCGTTGCCATTCGTCCTCCAAACCAAGGGCAGTT 1140
DB 1560 CATGTTTCCAGCGCAAAAGCTTGGCTCGGGCGTTGCCATTCGTCCTCCAAACCAAGGGCAGTT 1619
QY 1141 AGTTCCTCGGTGAGTGGAAAGATGTTGGTGGCATTCATCTGGCCATGCTTTCGCAAT 1200
DB 1620 AGTTCCTCGGTGAGTGGAAAGATGTTGGTGGCATTCATCTGGCCATGCTTTCGCAAT 1679
QY 1201 TCGCACCAAGCTGAGGATGGTTCCAAATGCTGATATCTTGATGACATTTGTTTGCACAC 1260
DB 1680 TCGCACCAAGCTGAGGATGGTTCCAAATGCTGATATCTTGATGACATTTGTTTGCACAC 1739
QY 1261 AGTAAACCTCAACCGCAGCAGCTTTAAACCCGCTGAAGACGAGGCGATGAAGTCAAAGC 1320
DB 1740 AGTAAACCTCAACCGCAGCAGCTTTAAACCCGCTGAAGACGAGGCGATGAAGTCAAAGC 1799
QY 1321 AGGGAGCTGCTGTGTGTAATTCGATTTGATGCCATTAAGGCTGCAAGTTATGAGTAAC 1380
DB 1800 AGGGAGCTGCTGTGTGTAATTCGATTTGATGCCATTAAGGCTGCAAGTTATGAGTAAC 1859
QY 1381 CACGCGGATTTGTTTTCGAATTCAGAAACCGGACCTGTAAACACTTACGTTTGGG 1440
DB 1860 CACGCGGATTTGTTTTCGAATTCAGAAACCGGACCTGTAAACACTTACGTTTGGG 1919
QY 1441 CGAAATTTGAAGCGGAGCCAACTGCTCAACGTGCAAAAGAAAGAGCGGTGCGCAAC 1500
DB 1920 CGAAATTTGAAGCGGAGCCAACTGCTCAACGTGCAAAAGAAAGAGCGGTGCGCAAC 1979
QY 1501 ACCA 1504
DB 1980 ACCA 1983

RESULT 8
BD093238
LOCUS DNA encoding sucrose PTS enzyme II. 5969 bp DNA linear PAT 27-AUG-2002
DEFINITION BD093238
ACCESSION BD093238
VERSION BD093238.1 GI:22638826
KEYWORDS WO 0102584-A/1.
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 5969)
AUTHORS Izui, M., Sugimoto, M., Nakamatsu, T. and Kurahashi, O.
TITLE DNA encoding sucrose PTS enzyme II
JOURNAL Patent: WO 0102584-A 1 11-JAN-2001.
AJINOMOTO CO INC, MASAKO IZUI, MASAKAZU SUGIMOTO, TSUYOSHI NAKAMATSU,

COMMENT	OSMU KURAHASHI	OS	Brevibacterium lactofermentum
	PN	WO	0102584-A/1
	PD	11-JAN-2001	
	PF	30-JUN-2000	WO 2000JP004348
	PR	02-JUL-1999	JP 99P 189512
	PI	MASAKO IZUI, MASAKAZU SUGIMOTO, TSUYOSHI NAKAMATSU, OSAMU FI	
	KURAHASHI		
	PC	C12N15/54, C12N9/12	
	CC		
	EH	Key	Location/Qualifiers
	FT	CDS	(3779)..(5761).
FEATURES		Location/Qualifiers	
source		1..5969	
		/organism="unidentified"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:32644"	
ORIGIN			
	Query Match	92.5%;	Score 1412.6; DB 6; Length 5969;
	Best Local Similarity	95.%;	Pred. No. 0;
	Matches 1462;	Conservative	0; Mismatches 64; Indels 1; Gaps 1;
QY	1	CTCATGSCATCTCGCGCGTTCTCGCGTTCCTGCCAGTGTGGTTGGTTTCACCGCAACCAAG	60
Db	4259	CTGATGGCATCTGCGCCGTTTCGGCGTTCTTGCCAGTGTGGTTGGTTTCACCGCAACCAAG	4318
QY	61	CGTTTCGGCGGCAATGAGTTTCCTGGGCGCGCGGATTGGTATGGCGATGGTTGCCGAG	120
Db	4319	CGTTTCGGTGGCAATGAGTTTCCTGGGCGCGG-GCAATTGGTATGGCGATGGTTCCCAAC	4377
QY	121	CTTTGGTGAACGGCTACGACGTGGCGCGCCACATAGCTGGCGGGGAAATGCCAATGTGGTC	180
Db	4378	CCTGGTTAACGGCTACGACGTGGCGCGCCACCAATGACCGCGGGGAAATGCCAATGTGGTC	4437
QY	181	CCGTGTTTGGTTTAGATGTTGCCAAGCGGTTTACCAGGACACCGTGTCTCTGTGCTGGT	240
Db	4438	CCGTGTTTGGTTTGGATGTTGCTCAAGCTGGTTTACCAGGGACCCGTGCTCTCTGTGCTGGT	4497
QY	241	GGTTTCTTGATTTCTTGGCAACGATCGAGAAGTTCTCTGCACAAGCGACTCAAGGGCACTGC	300
Db	4498	GGTCTCTTGATTTCTTGGCAACGATCGAGAAGTTCTCTGCACAAGCGACTCAAGGGCACTGC	4557
QY	301	AGACTTCCTGATCACTCCAGTGTGACGTTGCTGCTCACCGGATTCCTTACATTCAATCGC	360
Db	4558	AGACTTCCTGATCACTCCAGTGTGACTGCTGCTCACCGGTTCTTACGTTCAATTGC	4617
QY	361	CATTGGCCAGCAATGGCTGGGTGGCGGATGTCTGGGCACAGGTCACAGGCACTTTA	420
Db	4618	TATTGGTCCAGCAATGGCTGGGTGGGTGAGTCTTGTCTGGGCACAGGTCGAGGAGCTCTA	4677
QY	421	TGATTTCCGTTGGTCCAGTCCGGCGGTCTGCTCTTCGGTCTGCTACTCACCAATCCGTCAT	480
Db	4678	TGATTTCCGTTGGTCCAGTCCGGCGGTCTGCTCTTCGGTCTGCTACTCACCAATCCGTTAT	4737
QY	481	CATGTGFTGCACACAGTCTCTCCCGCCAAATTAGCTGGAGCTGTTTAAACAGAGGTGGATC	540
Db	4738	CACGTGFTGCACACAGTCTCTCCCGCCAAATTAGCTGGAGCTGTTTAAACAGAGGTGGATC	4797
QY	541	CTTTCATCTTCGCAACCGCATCTATGGCTATATATCGCCAGGTTGCGGCATGTTTGGCAGT	600
Db	4798	CTTTCATCTTCGCAACCGCATCTATGGCTATATATCGCCAGGTTGCGGCATGTTTGGCAGT	4857
QY	601	GTTCTTCTCGCGAAGAGTGAAGGCTCAAGGGCTTTCAGAGTGCCTTTCAGGTCCTCCGC	660
Db	4858	GTTCTTCTTAGCGAAGAGTGAAGGCTCAAGGGCTTTCAGAGTGCCTTTCAGGTCCTCCGC	4917
QY	661	TGTTCTTGGTATTACCGAGCCTCGGATCTTTCCGTTGTGAACCTTTCGCTGCGCTGGCGGTT	720
Db	4918	TGTTCTTGGTATTACAGAGCCTCGGATCTTTCGTTGTGAACCTTTCGCTGCGCTGGCGGTT	4977
QY	721	CTTTCATCCGTTATCGGTACCGCAGCTATCGGTGGCGCTTTGATTGCACTCTTTTAATATCAA	780


```
source 1. 1109
/organism="Corynebacterium glutamicum"
/mol_type="unassigned DNA"
/db_xref="taxon:1718"
<1. 1089
/feature="unassigned protein product; FRXA00315"
/codon_start=1
/transl_table=1
/protein_id="CAC27277.1"
/db_xref="GI:12579019"
/translation="YDFGPGVGLLFGLVYSPVITVGHQSPPIELFLNQGSFIF
ATANMAIAGGACLAIVFLAKSEKLGLAGSVSAVLGITEPAIFGVNLRWPF
IGTGAATAGGALIALFNKAVAGAGFLGVVSDAPDMVFLVCVVTFIAFGAAI
AYGLVLRNSISDPDAPAAVPVAGCTTKAABAPAEAFNSDSTLLOAPLTGEALSSV
SDAMPASGLSGVAIVTKGQLVSPVSGKIVVAPPSGHAFAVTRKAEDSGNVDLMH
IGFDVNLNTHFNPLKQGVDEKAGELLCEFDIDAIKAAAGTEVTPPIVSNYKKTGP
VNTYGLGIEZAGANLLNVAKEAVPATP"

ORIGIN
Query Match 72.5%; Score 1107.4; DB 6; Length 1109;
Best Local Similarity 99.9%; Pred. No. 4.8e-288; Indels 0; Gaps 0;
Matches 1108; Conservative 0; Mismatches 1;

QY 419 TATGATTTGGGTGTCAGTCCGGGTCTGCTCTTCGGTCTGGTCTACTCACAATCGTC 478
DB 1 TATGATTTGGGTGTCAGTCCGGGTCTGCTCTTCGGTCTGGTCTACTCACAATCGTC 60

QY 479 ATCACTGGTCTGACAGTCCTTCCGCCAATTGAGCTGGAGCTGTTTAAACAGGTGGA 538
DB 61 ATCACTGGTCTGACAGTCCTTCCGCCAATTGAGCTGGAGCTGTTTAAACAGGTGGA 120

QY 539 TCCTTCATCTTCGCAACGGCACTATGCTTAATATCGCCAGGCTGCGCANGTTGGCA 598
DB 121 TCCTTCATCTTCGCAACGGCACTATGCTTAATATCGCCAGGCTGCGCANGTTGGCA 180

QY 599 GTGTTCCTTCCTGCGAAGAGTGAAGAGCTCAAGGGCTTTGCAAGTGTCTTCAGTGTCTCC 658
DB 181 GTGTTCCTTCCTGCGAAGAGTGAAGAGCTCAAGGGCTTTGCAAGTGTCTTCAGTGTCTCC 240

QY 659 GCTGTTCTTGATATACGAGCTGCGATCTTCGGTGTGAACCTTCGCTGCGTGGCGG 718
DB 241 GCTGTTCTTGATATACGAGCTGCGATCTTCGGTGTGAACCTTCGCTGCGTGGCGG 300

QY 719 TTCTTCATCGTATCGGTACCGAGCTATCGGTGGCGCTTTCATTCACCTCTTAAATATC 778
DB 301 TTCTTCATCGTATCGGTACCGAGCTATCGGTGGCGCTTTCATTCACCTCTTAAATATC 360

QY 779 AAGGCAAGTTCGCTGGCGCTGCAAGTTCCTTGGGTGTTTCTTATTCATGCTCCAGAT 838
DB 361 AAGGCAAGTTCGCTGGCGCTGCAAGTTCCTTGGGTGTTTCTTATTCATGCTCCAGAT 420

QY 839 ATGTCATGTTCTTGGTGTGTCAGTGTACCTTCCTTCATGTCATTCGGCGCAGGAT 898
DB 421 ATGTCATGTTCTTGGTGTGTCAGTGTACCTTCCTTCATGTCATTCGGCGCAGGAT 480

QY 899 GCTTATGGCTTTACTTGTTCGCGCAACGCGCAATGATCCAGATCAACCGCTGCT 958
DB 481 GCTTATGGCTTTACTTGTTCGCGCAACGCGCAATGATCCAGATCAACCGCTGCT 540

QY 959 CCAAGTCCCTGCAGGAACGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1018
DB 541 CCAAGTCCCTGCAGGAACGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 600

QY 1019 TCCACCATCATCCAGGCACTTTGACGGTGAAGCTATTGCACTGAGCAGCGTCAGCAT 1078
DB 601 TCCACCATCATCCAGGCACTTTGACGGTGAAGCTATTGCACTGAGCAGCGTCAGCAT 660

QY 1079 GCCATGTTTCCAGCGGAAGCTTGGCTCGGGGGTGGCTTCCATCGTCCCAACCAAGGGCGAG 1138
DB 661 GCCATGTTTCCAGCGGAAGCTTGGCTCGGGGGTGGCTTCCATCGTCCCAACCAAGGGCGAG 720

QY 1139 TTAGTTCTTCGGTGTAGTGAAGAGTGTGGTGGCAATTCCTTCAGTCCGCAATGCTTCGCA 1198
```


Db	166112	AAAGTCGATCCAGCATTATCGTCTAGTGCTGTCTCATCTCTGGTTCGTCGCATTATTATTC	16605
Qy	892	AGCGATTGTTATGCC	907
Db	166052	AACCTTTATCTATGCC	166037
RESULT 11			
PDCRAFOPER/c	PDCRAFOPER	21838 bp	DNA linear BCT 23-MAY-1994
LOCUS	Pedococcus pentosaceus raffinose operon genes.		
DEFINITION	Pedococcus pentosaceus raffinose operon genes.		
ACCESSION	L32093		
VERSION	L32093.1 GI:493181		
KEYWORDS	agar gene; agas gene; agl gene; alpha-galactosidase; alpha-glucosidase; fructokinase; insertion element; permease; rafp gene; rafs gene; raffinose operon; regulatory protein; scrA gene; scrB gene; scrK gene; scrL gene; sucrose-6-phosphate; transport protein.		
SOURCE	Pedococcus pentosaceus		
ORGANISM	Pedococcus pentosaceus Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae; Pedococcus.		
REFERENCE	1 (bases 1 to 21838)		
AUTHORS	Leenhouts,K.J., Bolhuis,A.A., Kok,J.J. and Venema,G.G.		
TITLE	The sucrose and raffinose operons of Pedococcus pentosaceus PPEI.0		
JOURNAL	Unpublished (1994)		
COMMENT	On May 25, 1994 this sequence version replaced gi:475106. Original source text: Pedococcus pentosaceus (strain PPEI.0) DNA; Insertion sequence IS30 homolog (transposable element insertion sequence IS30 homolog, kingdom Prokaryotae) DNA; Pedococcus pentosaceus (strain PPEI.0) DNA; Insertion sequence IS3 homolog (transposable element insertion sequence IS3 homolog, kingdom Prokaryotae) DNA; and Pedococcus pentosaceus (strain PPEI.0) DNA.		
FEATURES	Location/Qualifiers		
source	1..21838 /organism="Pedococcus pentosaceus" /mol_type="genomic DNA" /strain="PPEI.0" /db_xref="taxon:1255"		
repeat_unit	487..510 /note="DR2; putative" /rpt_type=direct		
repeat_unit	488..561 /note="IR1; putative" /rpt_type=inverted		
repeat_unit	513..561 /note="DR1; putative" /rpt_type=direct		
terminator	2724..2760 /standard_name="terminator 4" /notes="putative"		
gene	complement(2929..3864) /gene="rafr"		
CDS	complement(2929..3762) /gene="rafr" /functions="regulation raffinose-operon" /note="putative" /codon_start=1 /transl_table=11 /product="regulatory protein" /protein_id="AA225562.1" /db_xref="GI:475107"		
-10_signal	complement(3837..3842) /gene="rafr"		
-35_signal	complement(3859..3864) /gene="rafr"		
gene	3909..5918		

Db 15264 TTGGTGCACCCCTATCTGGTGTCTACG-ATGGCATGATCATGGTCTTACCATCACT 15206
Qy 124 GGTGAACGGCTACGACGTGGCGCCACCATGGCTGGCGGCGAAATGCCAATGTGGTCCCT 183
Db 15205 AGTTAACGGCTATAGCGTTGCAACACCATGGCAGCGCGAAGATGGTTACTGGAAAGT 15146
Qy 184 GTTTGGTTAGATGTTGCCAAGCCGGTTACACAGGACCGTGTCTCTGCTGCTGGTGGT 243
Db 15145 CTTTGGGTTACAGTTGACAGCGCGGTATCAAGCCCAAGTCTCCAGATCTTAGGGCT 15086
Qy 244 TTCTTGGATTCTGGCAACGATCGAAGAGTTCTCTGCACAAAGGACTCAAGGGCAGCTGCAGA 303
Db 15085 CGCCTTCATTAGTACCCCTTGAAAAATCTTCCATAAACACATTAAGGGGCAATTGA 15026
Qy 304 CTTCTGATCACTCCAGTGTGACGTTGTGCTGTCCCGGATCTCCATTCATTCATGCCCAT 363
Db 15025 CTTACGTTTACCCCGATGTTTGCCATTGTGATTACTGGTTTCCCTTAACTTTTACAATCGT 14966
Qy 364 TGGCCACGAAATCGCTGGTGGCGATGCTGTGGCACACGCTCTACAGGACTTTATGA 423
Db 14965 TGGTCCCGTCTTACGACCGTGAGGATGATTAACGCTTAGGCTTATATAA 14906
Qy 424 TTTCGGTGTCCAGTGGCGGTCTGCTCTCTCGGTCTGGTCTACTCACCAGTCTCATCAC 483
Db 14905 CAGCACCGGCTGATTTGGTATGGGAATCTTTGGTTTATATATCTGCAATTTGTTATTAC 14846
Qy 484 TGGTGTGACACGATCTTCCCGCCCAATGAGCTGGAGCTGTT-----TAACCA 531
Db 14845 TGGTCTCCATCAAACTTTCCAGCAATCGAAACCCAGTTGTTGGCCAAATGTTGCCAAAAC 14786
Qy 532 GGGTGTGATCTTCACTTCGCAACGCACTATGCTATATATGCTATATCGCCAGGGTGGCGATG 591
Db 14785 TGGTGGTTCCTTATCTTCCAGTCCCTCGATGGCCAACTATTGGCAAGGTCGCGCCAC 14726
Qy 592 TTTCGAGTGTTCTTCTCGCGAAGAGTCAAAAGCTCAAGGCCCTTGCAGGTGCTTCAGG 651
Db 14725 TTTCAGTATTTCCTTGGCCACTAAGAGCCAAAGCAAAAGCCCTGACTCTTCAGCCGG 14666
Qy 652 GTCTCCCGTGTCTTGGTATTACGAGCTCGGATTCCTCGATGTCGTTGGAACCTTCGCTGG 711
Db 14665 GGTCTCAGCGTTACTCGGAATTACGGAACCTGCTATTTTCGGGGTCAACCTCAAGATGAA 14606
Qy 712 CTGGCGGTTCTTCATCGGTATCGGTACCGCATCGGTGCGGCTTTCATTCGACTCTT 771
Db 14605 ATTTCCATTGCTTTTGACGCAATTCCTTCAGGAATGCTTCAGCTTCTTAGGACTATT 14546
Qy 772 TAATATCAAGGAGTTCGTTGGCGCTGAGGTTTCTTGGGTGTTGTTTCTATTGATGC 831
Db 14545 CCACGTTCTGTCGTCGGCATGGTCCCGCATCAGTCATCGGATTCATTTTCGATTGCTTC 14486
Qy 832 TCCAGATATGGTCATGTTCTTGGTGTGTCAGTTGTTACTCTTCTTCATCGCATTCGCGC 891
Db 14485 AAGTCGATTCCAGCATTCATGCTCAGTGTGTCATCTCGTTCGTCGCAATTATTC 14426
Qy 892 AGCGATTGCTTATGC 907
Db 14425 AACTTTTATCTAATGCC 14410

PPSURFOP 21839 bp DNA linear BCT 24-MAY-1994
LOCUS P.pentosaceus (PPE1.0) sucrose and raffinose operons.
DEFINITION Z32771
ACCESSION Z32771.1 GI:493728
VERSION alpha-galactosidase; alpha-glucosidase; enzyme Iabc; fructokinase;
KEYWORDS insertion element; permease; regulator; sucrose.
SOURCE Pedicoccus pentosaceus
ORGANISM Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
Pedicoccus.
REFERENCE 1 (bases 1 to 21839)

AUTHORS Leenhouts,K.K.J., Bolhuis,A.A., Kok,J.J. and Venema,G.G.
TITLE The sucrose and raffinose operons of *Pedococcus pentosaceus* PPE1.0
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 21839)
AUTHORS Leenhouts,K.K.
TITLE Direct Submission
JOURNAL Submitted (27-APR-1994) Leenhouts K. K., Institute for Biological Sciences, Department of Genetics, Kercklaan 30, Haren, The Netherlands, 9751 NN
COMMENT On May 26, 1994 this sequence version replaced gi:475962.
FEATURES
 Location/Qualifiers
 1..21839
 /organism="Pedicoccus pentosaceus"
 /mol_type="genomic DNA"
 /strain="PPE1.0"
 /db_xref="taxon:1255"
 487..510
 /feature="DR2"
 /citation=[1]
 /rpt_type=DIRECT
 488..561
 /feature="IR1"
 /citation=[1]
 /rpt_type=INVERTED
 513..561
 /feature="DR1"
 /citation=[1]
 2724..2760
 /standard_name="terminator 4"
 /citation=[1]
 /feature="rafr"
 complement (2929..3864)
 /feature="rafr"
 complement (2929..3762)
 /feature="rafr"
 /function="regulation raffinose-operon"
 /standard_name="regulator"
 /citation=[1]
 /codon_start=1
 /transl_table=11
 /product="regulator"
 /protein_id="CAA83663.1"
 /db_xref="GI:475963"
 /db_xref="GOA:P43465"
 /db_xref="Swiss-Prot:P43465"
 /translation="MNGEYKTLANKSFBSNVLFQGEACLPNYTKGNVRDSVYIHY
 TOEGKGTAAAHNPATVLKAGDIFILPKGTFCFOADNDQPKYFWIGFSAGIRIEAM
 LGSGLAQCXVLRQYONGHYADLSLYKVLHNPNSLINDVLGSLIYRLFYDLLRWY
 PADATNKVSTEQENLAVSYLOENYSTGCTIMDLCHYLNLRSVLYTLFKTHANTSP
 QKLLKRLLEDAKRLSTSNNSVQSIANNVGVKDSFTFSKAFKRYSGASPSYRKSIG
 I"
 complement (3837..3842)
 /feature="rafr"
 /citation=[1]
 complement (3859..3864)
 /feature="rafr"
 /citation=[1]
 3909..5918
 /feature="rafr"
 3909..3914
 /feature="rafr"
 /citation=[1]
 3932..3937
 /feature="rafr"
 /citation=[1]
 3993..5918
 /feature="rafr"
 /function="raffinose transport protein"
 /standard_name="permease"
 /citation=[1]
 /codon_start=1
 /transl_table=11
 /product="permease"
 I"

```

/protein_id="CAA83664.1"
/db_xref="GI:475964"
/db_xref="GOA:P43466"
/db_xref="Swiss-Prot:P43466"
/translation="MQEENYKVGGRLLYGFCAKNDADYSLSGVLIIPITSHLFD
TGNKLDNRVSLVTLIMVIRIVELDFDFIGNAIDRTKNSPCHRPWVVGTVSS
IILLFLNGLGGLYAKNMIYLVFAILYITMDIFYSKDVGFWSMLPSLTDSRRE
KTATFARLGSTIGGGLVGLVPAVIFPSAKATSTGDNRFWFIPTALICILISAWG
VGLTRVDSIDIRKQDVTGVMEIPKALAKNDQLLAAVLAVPYGVGINLGSLEVY
YFTVINGKPKRSILSIINILFLGLIATLPVLVSKEFSKGVFAGCLVFMGLGIAFT
IAGSNLWLVLAATMGFPQOMVFLVLMVITDSVEYQKLGHRDESALSVRPLID
KFGGLSNGVVGQIATISGMTGTATASSITAAGQLHFKLTFAPFALMLLIATGIFSK
QPLTFEKGABIYAEIERWTRKFDNTDQAEKVTSLDLATPIAGQVPLAQVNDP
TFAAGTLGCGDFALPAPDTRQVFTTRHAGLVGNGVILVLIHIGLTVK
LRGTGTSYVEGQHQVQGDDELLEFWDPTFKQAGLDDVTIMVTYNSTFTMDWLVP
GOAVKATDNLQLHTKA"
5976..8177
/gene="agar"
CDS
5976..8177
/gene="agar"
/standard_name="alpha-galactosidase"
/citation=[1]
/codon_start=1
/evidence=experimental
/transl_table=11
/product="alpha-galactosidase"
/protein_id="CAA83665.1"
/db_xref="GI:475965"
/db_xref="GOA:P43467"
/db_xref="Swiss-Prot:P43467"
/translation="MSLITVDQANRVFHLNQLTSYFVAVEQGTLSHLYFGHVDHY
HGLRYPRVDRGFSNLPGSTDRTEFRDTPKEYSTAGEMDYLHPAAIVRHDTGANAL
YLVYQVRIEAGKPLSLGPAAEVDEBTEAETLIIVLQVSOVERFDLOVTVIRDRPV
VTRSVQVQNGHVAVNLKVASQIDFTDRQFETITLPGAHANERPEGSINYGIQT
FGSLRGTSHQMPFLALVDHTTETSEDAYGNLVYSGNHAFELKQDQDLHLMWG
INSINFWKAGATQTEPVLVMVYTNKGLNAMSQAYHHLIRRVRSFBNQERPIV
VNWAEATFDENFAKLPIDEAKQLGIEMFVLDGDFGRDDNSLGDWQVDRKPF
PQGLNHFYVHQHFGKGLIPEMISYDSKLYQOHPDYLMOVPGKSPSPSRNOYL
DLGQAVRNRIFDQLDLKSKQIDYIKWMNRHLSDIYSVALPPERQGVVHRVYLG
LYELLERLTATYPHILFEGCGGGGFADGMAVYMQIWASTDVAERITIQYGISL
APYLSIATHSVSVSNQOQGRETSRNASAGSVGLGELDLTQLSSADKQIVQKQV
VQKQLRPLIQFEFRVLSPIITSNOAAWVFSPQDEAIVFVFNLTSTVAQPSLRTK
LVGLNPKLNYQNIATKAIFFGDELQMGFDYVPVYDYTTKVYHFAVTEN"
8265..8289
/standard_name="terminator 3"
/citation=[1]
9403..9426
/notes="DR2"
/citation=[1]
/rpt_type=DIRECT
complement(9519..10424)
/gene="iso-IS30"
complement(9519..10424)
/gene="iso-IS30"
/function="IS-30 (E.coli) like IS element"
/standard_name="IS-element"
/citation=[1]
/codon_start=1
/transl_table=11
/product="insertion element"
/protein_id="CAA83666.1"
/db_xref="GI:475966"
/db_xref="GOA:Q52261"
/db_xref="TrEMBL:Q52261"
/translation="SDAVTKRSASLKWCLPQTHLLGADLLSITYSERIKIETPCEL
GLSNQMGVLRNRPSTIYSLSRCQPYQAEIAQTDABYKRSQCGRKTLKSLDKQKI
LNLHLSWSPHNFQFKLATKSIYNLQNGRIGFSLNDLPEHGVQRNVRQSRKNY
QSLGISIEQRPMFNRIGNIDFELDTVVGPRGHSKAVLLTIDRKSRLWAYRLKDR
TTATVNEALTKEFLTTPNGPVHSTVDNRGTEFSGVLSLSBQYKITYCHAYTPADVW
NNALGIYVVFILKGLVLSLVLIKI"
10396..10442
/notes="DR1"
/citation=[1]

```

```

/rpt_type=DIRECT
10397..10443
/notes="IR2"
/citation=[1]
/rpt_type=INVERTED
10681..10716
/standard_name="terminator 2"
/citation=[1]
complement(10727..11593)
/gene="scrk"
complement(10727..11593)
/gene="scrk"
/function="fructose 6-phosphotransferase"
/standard_name="fructokinase"
/citation=[1]
/codon_start=1
/transl_table=11
/product="fructokinase"
/protein_id="CAA83667.1"
/db_xref="GI:475967"
/db_xref="GOA:P43468"
/db_xref="Swiss-Prot:P43468"
/translation="MLLGAIBAGTKFVCATGAENGQVSDRISIPITTPVETMTAVDD
YFTHEPDAIGIGSFGPIGVNPHDPKYGIYITTPKPGWGFDFLGHKLSQFNIPLYWT
TVNEAAYGESIMIGIAKVPNSIYMTIGTVGAGVISOHIFNGRTHTELGHWRLNL
PGDDFKSNCPYHDICLEGAAAGPAVKRTGKAGKDI PVDDPWPIITDYIAQACVNL
VAPAPDKIILGGVVMNQRLFPPIREKFPAYLNGYEEVPLDDIIVPAGLNNNSGIAG
GLLLAQAALKNA"
complement(13852..15989)
/gene="scrA"
complement(13852..15989)
/gene="scrA"
/function="enzyme II of the PTS system, sucrose specific"
/standard_name="enzyme IIscr"
/citation=[1]
/codon_start=1
/evidence=experimental
/transl_table=11
/product="enzyme IIabc"
/protein_id="CAA83668.1"
/db_xref="GI:475968"
/db_xref="GOA:P43470"
/db_xref="Swiss-Prot:P43470"
Query Match 18.8%; Score 287.6; DB 1; Length 21839;
Best Local Similarity 58.8%; Pred. No. 4.3e-66;
Matches 539; Conservative 0; Mismatches 364; Indels 13; Gaps 2;
QY 4 ATGGCATCTGCGCGCTTCGCGTTCTTCCAGATGTTGGTTGGTTTCCACCGCAACCAAGCGT 63
Db 15324 ATGGCTAGTGGCGCGTTTACTTCTTCCCAATCCTCTTAGGATTTTCAGCAAGAGCGC 15265
QY 64 TTGCGCGCAATAGTTCCTTGGCGCGCGGTATGTGTATGGCGATGTTCTCCGAGCTT 123
Db 15264 TTGCGTGGCAACCCCTATCTGGTGCTACG-ATGGCGATGATCATGTTCTTACCATCCT 15206
QY 124 GGTGAACGCTACGACCTGCGCGCCACCATGGCTGCGGCGGAAATGCCAATGTGTCCT 183
Db 15205 AGTTAACCGCTATAGCGTTGCAACGACCATGGCAGCGCGCAAGATGTTTACTGGAACGT 15146
QY 184 GTTTGGTTTAGATGTTGCCAAGCGGTTTACAGAGCGCAACCGTCTTCCTGCTGCTGCTGT 243
Db 15145 CTTTGGTTTACGTTGCAAGCGCGCTATCAAGCGCAAGTGTCTCCAGTCTTAGGCGT 15086
QY 244 TTCTTGGATTTGGCAACGATCGAGAAGTTCCTGCAACAGCGCTCAAGGGCACTGCAGA 303
Db 15085 CGCCTTCACTTAGCTACCTTGGAAAAATCTTCCATAAACAATTAAGGGGCACTTCA 15026
QY 304 CTTTCTGATCATCTCAGTCTCAAGTGTGCTGCTACCGGATTCCTTCAATTATGCCCAT 363
Db 15025 CTTACGTTTACCGCGATGTTTGCATTTGATTAATGTTTCTTAACTTTTACAATCGT 14966
QY 364 TGGCCAGCAATGCGCTGGTGGCGGATGTGCTGGCACACGCTCTACAGGACTTTATCA 423

```

terminator

repeat_unit

gene

CDS

repeat_unit

Db	14965	TGGTCCCGTCTTACGACGGTGAGCGATGCAATTAACGGTTAGTAGGCTTATATAA	14906
Qy	424	TTTCGGTGGTCCAGTCGGCGGTCTGCTCTTCGGTCTGTCTACTACCAATCGTCATCAC	483
Db	14905	CAGCACCGGCTGGATGCTATGGGAATCTTTGGTTTATATATCTGCAATGTTATTAC	14846
Qy	484	TGGTCTGACAGCTCTTCCCGCCAAATGAGCTGGAGCTGTT-----TAACCA	531
Db	14845	TGGTCTCCATCAAACTTTCCGAGCAATCGAACCCAGTTGTGGCCATGTTGCCAAAC	14786
Qy	532	GGGTGGATCTTCATCTTCGCAACGGCATCTATGGCTAATATCGCCAGGGTGGCGCATG	591
Db	14785	TGGTGGTCTCTTATCTTCCAGTCGCTCGATGGCCCAACATTTGGGCAAGTGGCGCCAC	14726
Qy	592	TTGGCAGTGTCTTCCTGGCGAAGAGTGAAGAGTCAAGGCCCTTCAGGTGCTTCAGG	651
Db	14725	TTTAGCTATTTCTTTGGCCACTTAAGAGCCAAAGCAAAAGCCCTGACTTCTTCAGCCGG	14666
Qy	652	TGCTCCGCTGTCTTGGTATTACGAGGCTCGGATCTTCGGTGTGAACCTTCGCTCGG	711
Db	14665	GGTCTCAGGTTACTCGGAATACGGAACCTGCTATTTTCGGGGTCAACCTCAAGATGAA	14606
Qy	712	CTGGCGGTTCTTCATCGGTATCGGTACCGCAGCTATCGGTGCGGCTTTGATGCACTCTT	771
Db	14605	ATTTCCTATTTGCTTTGACAGGATTTGCCCTCAGGAATGCTTCAGCTTTCTTAGGACTATT	14546
Qy	772	TAAATCAAGGAGTGTGGTGGCGCTGCAGGTTCTTCGGTGTGTTCTTATTAATGATGC	831
Db	14545	CCACGTTCTGCTGGCGATGGGTCGCCCATCAGTCATCGGATTCATTTGCAATGCTTC	14486
Qy	832	TCCAGATATGTCATGTTCTTGGTGTGTCAGTGTGTAACCTTCTTCATCGCATTCGGCGC	891
Db	14485	AAAGTCGATTCAGCAATCACTCACTGCTGCTCATCTCGTTCGTCGCTGCAATTTATCC	14426
Qy	892	AGCGATTTGCTTATGGC 907	
Db	14425	AACTTTATCTATGCC 14410	
RESULT 13	AF401046	2788 bp DNA linear BCT 18-MAR-2002	
LOCUS		Lactobacillus sakei sucrose-specific enzyme II of the pts (scra)	
DEFINITION		and dextran glucosidase (dexB) genes, partial cds.	
ACCESSION	AF401046		
VERSION	AF401046.1	GI:15217137	
KEYWORDS			
SOURCE			
ORGANISM		Lactobacillus sakei	
REFERENCE		Lactobacillus sakei	
AUTHORS		Bacteria: Firmicutes; Lactobacillales; Lactobacillaceae;	
TITLE		Lactobacillus.	
JOURNAL		1 (bases 1 to 2788)	
MEDLINE		Dudez, A.-M., Chaillou, S., Hissler, L., Stentz, R.,	
PUBMED		Champomier-Verges, M.C., Albert, C.A. and Zagorec, M.	
REFERENCE		Physical and genetic map of the Lactobacillus sakei 23K chromosome	
AUTHORS		Microbiology 148 (Pt 2), 421-431 (2002)	
TITLE		21822493	
JOURNAL		11832506	
MEDLINE		2 (bases 1 to 2788)	
PUBMED		Dudez, A.-M., Chaillou, S. and Zagorec, M.	
REFERENCE		Direct Submission	
AUTHORS		Submitted (20-JUL-2001) Flore Lactique et Environnement Carne,	
TITLE		INRA, Domaine de Vilvert, Jouy-en-Josas 78350, France	
JOURNAL		Location/Qualifiers	
FEATURES		1..2788	
source		/organism="Lactobacillus sakei"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:1599"	
gene		<1..1605	
		/gene="scra"	
CDS		<1..1605	
		/gene="scra"	

		/note="scra; similar to Scra of Pedicoccus pentosaceus"	
		/codon_start=1	
		/transl_table=11	
		/product="sucrose-specific enzyme II of the PTS"	
		/protein_id="AAK92528.1"	
		/db_xref="GI:15217138"	
		/translation="FVPLIPALTAGGLLMAINNVLTGQGLFGAOSIVOMFPQWKGFAE	
		IVNMSSAPFTFLPILAFSAKTFGGNYPYLGAAAGMMLVMPNLVNGYVAESATIGH	
		MTYWHVGLNIAQAGYGOVPIVIGVAFILANBKFFHKLHNDVDFTFPMLSIIT	
		GFLFTLVGPALRIQVSVGTSLVAVYOTLGAVMGIFGLGYSAILVTLGHOSPPAIE	
		TLLIADAKTGGSPFIPVAAAMANTAGGAATEAVFTVKNKOOKSLTTSAGISAPMLGPT	
		EPALFVGNLKURFPFFGLIASGSSFFIIGLHLVLSVWGPAGIIGFIALAPKISPT	
		MMGAIISFVIAFVGTLYKKAMKTETEEIINEAPATPEVVEVDEKISAPVTRIV	
		DLASVPDFVASEAMGKIATMPTSDVLAFTVGTIVTAAANTHAYGKSDGDAVLI	
		HIGLDTVNLNGIGPEKIVQOQHVSEGLLGHFDIDIKIQAAGLPLTMTVTNTAGYA	
		OVDPLLTVDKAAMQGEELIQLHAKKD"	
gene		1604..>2788	
RBS		/gene="dexB"	
		1604..1610	
		/gene="dexB"	
		1617..>2788	
CDS		/gene="dexB"	
		/note="DexB; similar to DexB of Streptococcus mutans"	
		/codon_start=1	
		/transl_table=11	
		/product="dextran glucosidase"	
		/protein_id="AAK92529.1"	
		/db_xref="GI:15217139"	
		/translation="MOTNMWNAFVQVYPRSFDSNGDIGIIOIIRLDYLDLADLG	
		VNAIWSVYSPNVNDGYDISDQAINPEYGSVMVDSQLLEAKIRIKIWMIDLAVN	
		HTSDQHPLEARKSKONPFDYIWRDPAIDGDSVPNDLQSNFKSGNAFADVATQYY	
		LHFYAKEQFDLWQNPVREAVYQMTWLQKIGGFRMDVLDLIGKPEFDKIKENGP	
		OLHAYIQEMNARVLSQYDVTVVGETWGATPEIQMYSNPRHLSMIQFEGINIDKQ	
		SGMTWRDLKPLIPAEHLAVFVKQLADGVGNLSLFWSNHDLPRIVSRWGDSDQVREK	
		SAQALAIYHMLKTPYIQGEIEGNTVPIITSYHEIEDIESRRVYQROQQQYAI"	
ORIGIN			
		Query Match 16.6%; Score 253.4; DB 1; Length 2788;	
		Best Local Similarity 51.9%; Pred. No. 6e-57;	
		Matches 734; Conservative 0; Mismatches 621; Indels 58; Gaps 5;	
Qy	2	TCATGGCATCTGGCGGTTCGGGTCTTCCAGTGTTCGGTTCACCGCAACCAAGC	61
Db	143	TCATGTCTGTCACGGCAATTACCTTCTACCAATTTAATTCATTTTCGCACTAAGC	202
Qy	62	GTTTCGGCGGCAATGAGTTCTTCGGCGCGCGCTATTTGTTATGCGATGTTTCCGAGC	121
Db	203	GATTTGGTGGAAATCCATATTTAGTTCGGCGAGCA-GGAATGATGTTGTTATGCTTAAC	261
Qy	122	TTGGTGAACGGCTACGACGTGCCGCCACCATGGTTCGGCGGCAATGCAATGTTGCTTC	181
Db	262	CTAGTTAATGTTTATGGGTTCGTGAAAGTATCGCTACAGGCGATATGACGTATTCGCAT	321
Qy	182	CTGTTTGGTTTATGATGTTCCCAAGCCGTTACCGGCGCACCGTCTCTGTCGTGGTG	241
Db	322	GCTTTGGATGCAATATTCGCAAGCTGTTATCAAGGTCAAGTAAATTCGGTTATCGGA	381
Qy	242	GTTTCTTGGATTCCTGGCAACGATCGAGAAAGTTCCTGCACAAAGCACTCAAGGCACTGCA	301
Db	382	GTTGCCTTCATTTCTTGCTAATTTAGAGAAATTTCTCCATAAATTAATGATGCTGTT	441
Qy	302	GACTTCCTGATCACTCCAGTGTGAGCTGCTCACCAGGATTCCTTCATTCATTCATCGCC	361
Db	442	GATTTTACGCTTTACACCAATGTTATCGATTATATCACAGGATTTTAAACCTTACATTA	501
Qy	362	ATTGGCCACGACATCGCTGGGTGGCGGATGTCGGCACACCGTCTACAGGACATTTAT	421
Db	502	GTTGGTCTGCAATGAGAAATGTTTCAAAAGGTGTTACTGATAGTTTAGTTGGGCTTAT	561
Qy	422	GATTCGGTGGTCCAGTCCGCGGTCTGCTCTTCGGTCTGCTCTACTACCAATTCGTCATC	481
Db	562	CRAACGTTAGTGACGTGGTATGGGATTTTTCGGCTGGGTTATTTCAGCAATTGCTTA	621

RBS	/gene="sacB"		3426	TTTGGCGGAAATCCATATTTAGTGCTG- TTGTAGGTATGATTAGTGCCTGGACT	3484
	/evidence=experimental				
	complement (2721. .2725)				
	/gene="sacA"		124	GGTGAACGGCTACGACGTGGCGCCACCATGGCTGGCGGCGAAATGCCAATGTGGTCCCT	183
	complement (2737. .2743)				
-10_signal	/gene="sacA"		3485	TATTAAATGGATATATATGTCTGAAGCAATCTTAATCATCAATAGATTATTTGGGATAT	3544
	/evidence=experimental				
	complement (2761. .2767)				
	/gene="sacA"		184	GTTTGGTTTATAGATTGGCCAAAGCGGTTACAGGGCAACCGTGCTTCCTGTGCTGGTGT	243
	/evidence=experimental				
gene	2815. .4835		3545	TTTTGGTTTTAAAGTTGCCAGGCTGGATATCAAGGACAAGTCTCTCTGTAATTGGAGT	3604
	/gene="sacB"				
	2815. .2821		244	TTCTTGGATTCTGGCAACGATCGAGAGTTCTCTGCAAGGAGCTCAAGGCGACTGCAGA	303
	/gene="sacB"				
	/evidence=experimental		3605	GGCTTTTATCTCTGCTAAACTTGAGAGATTCTTTTATATAACCTTAAACGATGCTATAGA	3664
-35_signal	2837. .2843		304	CTTCTGATCATCTCCAGTGTGAGTGTCTCTCACCGGATTTCTTACATTCATCGCCAT	363
	/gene="sacB"				
	/evidence=experimental		3665	CTTCACATTTACTCCGTTACTTTTCTGATCATCACTGATGTTCTCAGATTCATCTATGT	3724
	2872. .2877				
	/gene="sacB"		364	TGGCCAGCAATGGCTGGGTGGCGATGTCTGGCACACGGTCTTACAGGGAATTTATGA	423
RBS	2883. .4835		3725	GGGACCACTCTCGTTTGTATCAAAATGGCTTGACAGATGGTTGGTTGGACTTTATAA	3784
	/gene="sacB"				
	/EC_number="2.7.1.69"		424	TTTCGGTGTCCAGTCTGGCGGTCTCTCTTGGTCTGGTCTACTCACCAGTCTCATCAC	483
	/codon_start=1				
	/transl_table=11		3785	CACATTAGGGGCACTAGGAATGCTTGTCTTGGAGGCTTCTACTCGGCAATAGTAGTAAC	3844
CDS	/product="enzyme II sucrose protein"				
	/protein_id="CAB09690.1"		484	TGGTCTGCACCACTCTTCCGCCCAATTTGAGTGGAGCTGTTTAAACCG- -----	532
	/db_xref="GI:4581477"		3845	TGGATTACATCAGAGTTTTCGGCCATTTGAACAATGTTAATCAACAACATCATCAACAG	3904
	/db_xref="GOA:Q9S6S6"				
	/translation="NNHKQVABRIILNAVGRDNIQGRHCAATRLRLVLKDTGVIDQEAL		533	-----GGTGGATCCCTCATCTTCGCAACGGCACTATATGGCTAATATCGCCAGGGTGC	585
gene	DNDPLKGTFFRAAGQYQIIVGPDVNTVVEEPIKLTGISEASTADLKFIAGSQKQNP		3905	TGGTATTGGTGGAGACTTTATTTTCCAGTTGCGGCTCGCAAAATATGGCTCAAGCTGG	3964
	VMAVKLSDLPVPLIPALVAGLLMALNNALTAHLFATKSLVEMFPMWKGPAIDVN				
	TMSAAPTEMPILLICYSATKPGNPYLGAVVGMIMVMPGLINGVVAEAI SNHWY		586	GGCATGTTTGGCAGTGTCTTCTCGGCGAAGATGAAAAGCTCAAGGGCTTGCAGGGTGC	645
	WDIFKVAQAGQGVQVPIGVAILAKIRFPFKILNDALDFTFELLSVIINGFL		3965	TGCACTTTTTCGAAATTTATTTTGTACTAAAAACAATAAGACAAAGCTCTTGCAGCTCC	4024
	TFTIIVGPAIRFVNSGLTDGLVNTLGAJLNVFGFYSAIVVTLHQSPFPAITML		646	TTTCAGTGTCTCCGCTGTCTTGTGTATTACGGAGGCTCGCATCTTTCGGTGTCAACCTTCG	705
RBS	ITNYQSGTGGDFIPFAACANMAQAGATFALLFTVKNTKTKALAAPAGVSAIILITE		4025	AGCTGGTGATCTGCTATTCTTAGGTATTACAGAACCTCGGTATTTTGGGATTAATCTAA	4084
	PALGINLKVPFTIAGASAIGSLFMGLFHLAVSLGSLIGFISTIKAGINLQFM				
	TSIFSLIAKVPFTSIYGRMEPAKSTKEKNKONATTOYOPBKVIIDPVKSGELLAPI		706	CCTGCGCTGGCGGTTCTTCATCGGTATCGGTACCGAGCTATCGGTGGCGCTTTGATTCG	765
	NGVPIPLSDVDSPVSKIMGMGIALKPKSGELFSPAGEIILIAVETGHAYIKTNG		4085	ACTAAAATATCCGTTCTTTTATTTGCTTTGGGCTTTCAGCAA TTGGTTCAATTTATGGG	4144
	GEVLHIGDITVSMNGFIQNVKVGQVKAGDGLLGSFDEKBEIKKSGLDLDIVLITIN		766	ACTTTTAAATATCAAGGCAAGTTGCGGTGGCGCTCGCAGGTTTCTTGGGTGTTGTTCTAT	825
CDS	SASYNIEILSENVDIKVGEKILLN"		4145	ATTAATCCATGTCTTGGCGTTAGTCTGGGATCCGAGGAATTAATTTGGCTTTTATCTCTAT	4204
	4838. .5720				
	/gene="sack"		826	TGATGCTCCAGATATGTCATGTTCTTGTGTGTGTCAGTTGTTACTCTTCTCATCCCAT	885
	4838. .4842				
	/gene="sack"		4205	TAAAGCTGGGTATAACTTACAAATTAATTAATTTGATTTATTTAGCTTTCTTATTGGCTT	4264
gene	4848. .5720				
	/EC_number="2.7.1.4"		886	CGCGCAGCGAATGTCTTATGGCTTTTACTTGGTTCCGCCAACCGCAGCATTTGATCCAGA	945
	/codon_start=1		4265	TGTTGTACTCAATCTATGTCGGCGATGGAAGCTAAATCTATTACGAAAAGAAAAA	4324
	/transl_table=11				
	/product="fructokinase"		946	TGCAACCGCTGCTCCAGTCCCTGCGAGAACGACCAAGCCGAGCAGAGACGCCGACAGA	1005
RBS	/protein_id="CAB09691.1"		4325	TAAACAGAAATGC-----AACAACTCAATACCAACCTCGAGAAAAGTTATTAT	4369
	/db_xref="GI:4581478"				
	/db_xref="GOA:P82371"		1006	ATTTTCAACAGATTCACCATCATCCAGGCACCTTTGACCGGTGAAGCTATTTCGACTGAG	1065
	/translation="MSVYVYGEAGTKFVLAIADBEHNIKKKFPATTTPOETISKT		4370	CGATCCAGTTAAAGTGGCGAACTCTTGTCTCGATAATGATTTGTGATTTCCCTCTGTC	4429
	IKYFENRVAISALGSRFDINLSKTYGTTSTPQVKNINLVGLKEALDIPY		1066	CAGCGTCAGCGATGCCAATGTTTTCAGCGGAAAGCTTGTGCTCGGGGTTTGCATGCTGCC	1125
CDS	FTFDNVAAGEMKNIGKNLVLTITGTGGGALQNGYFPGIGHSEMGHORINRHR				
	DLVTFEGTCPPFGDCLGVAGAPSLFARTGILGEKISSDDPIWDILSLVYIAQAAINAT		4430	TGATGTAAAGTGAACCTCTGTTTCTCAAAAGAAAATATGGGAAAAGGTTATTTCGAATCAG	4489
	LTLAPECIILGGVMEKPNMISLIQKFIISMLNVIDLPCSVKVIIRLPTVKENGSA				
	LGNFLAYLSLFTKE"		1126	AACCAAGGGGCGATTAGTTTCTCCCGTGAGTGGAAAGATTGTTGGTGGCATTCCTCATCTGG	1185
	ORIGIN				
Query Match		16.4%; Score 250.8; DB 1; Length 5800;			
Best Local Similarity		51.1%; Pred. No. 3.3e-56;			
Matches 725; Conservative		0; Mismatches 647; Indels 46; Gaps 4;			
QY	4	ATGGCATCTGCGCGTTTCCGCTTCTTCCAGTGTCTTGGTTGGTTTCACCGCAACCAAGCGT	63		
Db	3366	ATGAGTGGCGCCACCATTTACTTTATGCCAAATTTTAAATAGTTATTCTGCCACAAAAGA	3425		
QY	64	TTTCGGCGGCAATGAGTTCTCTGGCGCGCGGTATTGGTATGGCGATGGTGTCCCGAGCTT	123		

```

Db 4490 TAAATCGGAGAACTTTTCTCACCAGCCGAGGAGAAATTAATTATTCGTTATGAAACAGG 4549
1186 CCAATGCTTTTCGAGTTCGCACCAAGGCTGAGGATGTTCCAAATGTCGATATCTTGATGCA 1245
Db 4550 TCACGCTTATGTTATTAACAAACAAATGAGGT-----GAGTCTTTTACA 4597
QY 1246 CATTGTTTCGACACAGTAACTCAACGCGACGCTTTTAAACCGCTGAGGAGCAGGG 1305
Db 4598 TATTGGAATAGATACCTGTTTCAATGAATGTAATGATTCATACAAAATGTTAAAGTTGG 4657
QY 1306 CGATGAAGTCAACAGCGAGGAGCTGCTGTGTAATTCGATATTCGATCCCATTAAGGCTGC 1365
Db 4658 CCGAAAGTAAACGACGAGGAGTCTTCTAGCATCTTTGTAAGAAGAAATCAAGAAGAG 4717
QY 1366 AGCTTATGAGTAAACACGCGCGGATGTTGTTTCGAAT 1403
Db 4718 TGAATGGAATGACTGTAAATTAATTGTTATTACAAAT 4755

RESULT 15
AE004395/c
LOCUS
DEFINITION Vibrio cholerae O1 biovar eltor str. N16961 BCT 10-OCT-2003
52 of 93 of the complete chromosome.
ACCESSION AE004395
VERSION AE003853
KEYWORDS
SOURCE
ORGANISM
Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
Vibrionaceae; Vibrio.
Vibrio cholerae O1 biovar eltor str. N16961
Vibrio cholerae O1 biovar eltor str. N16961
Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
Vibrionaceae; Vibrio.
1 (bases 1 to 10264)
Heidelberg, J.F., Eisen, J.A., Nelson, W.C., Clayton, R.A., Gwinn, M.L.,
Dodson, R.J., Haft, D.H., Hickey, E.K., Peterson, J.D., Umayam, L.,
Gill, S.R., Nelson, K.E., Read, T.D., Tettelin, H., Richardson, D.,
Ermolaeva, M.D., Vamathevan, J., Bass, S., Qin, H., Dragoi, I.,
Sellers, P., McDonald, L., Utterback, T., Fleischmann, R.D.,
Nierman, W.C., White, O., Salzberg, S.L., Smith, H.O., Colwell, R.R.,
Mekalanos, J.J., Venter, J.C. and Fraser, C.M.
DNA sequence of both chromosomes of the cholera pathogen Vibrio
cholerae
Nature 406 (6795), 477-483 (2000)
20406833
10952301
2 (bases 1 to 10264)
Heidelberg, J.F., Eisen, J.A., Nelson, W.C., Clayton, R.A., Gwinn, M.L.,
Dodson, R.J., Haft, D.H., Hickey, E.K., Peterson, J.D., Umayam, L.A.,
Gill, S.R., Nelson, K.E., Read, T.D., Tettelin, H., Richardson, D.,
Ermolaeva, M.D., Vamathevan, J., Bass, S., Qin, H., Dragoi, I.,
Sellers, P., McDonald, L., Utterback, T., Fleischmann, R.D.,
Nierman, W.C., White, O., Salzberg, S.L., Smith, H.O., Colwell, R.R.,
Mekalanos, J.J., Venter, J.C. and Fraser, C.M.
Direct Submission
Submitted (14-JUN-2000) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA
FEATURES
Location/Qualifiers
1..10264
/organism="Vibrio cholerae O1 biovar eltor str. N16961"
/mol_type="genomic DNA"
/strain="N16961"
/serotype="O1"
/db_xref="taxon:243277"
/note="biotype: El Tor"
/chromosome="II"
complement (114..1553)
/gene="VCA0653"
complement (114..1553)
/gene="VCA0653"
/note="similar to GB:M76768 SP:P22825 PID:155262;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/protein_id="AAF96554.1"

```

```

/db_xref="GI:9658069"
/translation="MDYPVIAKQLLESLOGKSNIQALAHCACTRLRLVLDNTOINSEA
ISLQVVKQFKVAGYQIIFGSGIVNQVIAEMAKLTGLVEMSTNDVASAGAKQNSWA
QRAVGLSDIFVRIIPAI VAGGLMGLFNVLTPAGLFIQESLIDNPGLADLASMIN
TTPANFPVILPVLAFASASRKTGPNFLGAALGMLMVPDLNGLNGWFGSGASVGTPT
WNLAFIEIKVGYQGSVLPIVNSAYILAKIENGLRIKTVPSVDNLLTLMIAITGFL
TFTVGPLTRDVGEMLDALNMLYDSAGVGGALFGFIYAPFVITGMHHSFIALEIOL
LADIVTGTGTFIIPFAAMSNIAGGAALAVGVMTKETKLKGVAIPSGVTALLGITEPA
MGVNLKLRYPPIAICGAALASAFITLFFNVKAQALGAAGLPGIISINPOQIYYING
MAISFVAAFALIVLLAMREKTKQAQVTA"
1854..2846
/gene="VCA0654"
1854..2846
/gene="VCA0654"
/notes="similar to GB:114685 PID:289864; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="sucrose operon repressor ScrR, putative"
/protein_id="AAF96555.1"
/db_xref="GI:9658070"
/translation="WASLHDVARLAGVSKTSVSRVINDEYGVKESKIKVLKAVEQCG
YMWVAKDLKSKQTNLIGVIVRVSSHATAQGVGLTATFEQAGKHVLIANTHQVHA
KEIETQIFNQRRVEGIVFYATHLQPLVHAHQHSAVPPVVLVQDQSLNIPSIHDD
HRVFEAGQRLVKAGCKQIFIGVQSDDIADVLRSGQALQALGPHQOITLPHARGNF
SIESGYOMAKOAVLDYIPOLDGLFCATDRLAVGAIKALQELGVHVGQVLLGVGNDEL
AVVSNPPISTFNYPADKAGENAAMKLLERIAGRQEMSKVILTFQNIQRETCPV"
2322..4568
/gene="VCA0655"
2322..4568
/gene="VCA0655"
/notes="similar to GB:M74035 PID:155269; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="sucrose-6-phosphate dehydrogenase"
/protein_id="AAF96556.1"
/db_xref="GI:9658071"
/translation="MLDTLLLAGINNVTILAPQGVVLAKHPPLVPHLPDDVS
LOSILGQWLSVQRTAEVSDQOLAAGKAIAEROKLETLPYQTDLCPYRLWHISPP
QGLNDPNGFIYHGEYHLFYOMHFACEHDKYVWHLKSLDLVDWQWQSVALTSDW
FDSHGVSCHAVSHQOQDLMLFYTNTRLCVDRQRTQCAARMWANGPEKLGPIVRC
LPEGVTEHIRDPKVITQCKWMLGAOTLAHQGLIYVVSDDLHWHFDPKLYGDELG
DYGWMECPDMFELOGEAFFVPGQGLASANPHHTIEHQRIFRATQNAQGLIALQG
WPLDGFPPYADQATQADGRVLCGMGLPDETQHPSCDQWQIHLQTLALRELEWKEG
RIYQPLRELDLQSEPHLTLSDNVTELKTSFALQVTLPGWCELRMLMONTQYRVL
TIDAEQLRLDRSATQIQGDTTIRELKLDSPTVELRLADQSSLEIFINQGHVMTS
RIPTPDASGISLHGASVDKLYMAPASAPTNLEVNVP"
complement (4690..5661)
/gene="VCA0656"
complement (4690..5661)
/gene="VCA0656"
/notes="similar to GB:M76768 SP:P22824 PID:155263;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="fructokinase"
/protein_id="AAF96557.1"
/db_xref="GI:9658072"
/translation="MNALVRLSSNHLRSDSMSRVLTGDAVVDLIPDQOQHYLKCPG
GAPANVAIARLCGRSAFFGRVGNDFRGMQOQLTLDQVDCOHLVDPVHRTSTVV
VDLDEHGRSFTFMVKPSADQFLQSLDIPSQNGEWLHVCSIALANQSRSTPAALIA
QMKEVGVSFEDPNLREVRVEFQELQAVIMRAVGLADVRFSEBELQFLTGTSIEE
GLQIADPQIPLVVVTLGAKGALVVTNSRQVAVKVPDITTTGAGAFVGGLLYRL
SVAQDHNQATILDVAKWANGCAGALATTQKAMTALPNQAALYAFLE"
6008..7546
/gene="VCA0657"
6008..7546
/gene="VCA0657"
/notes="similar to SP:P13035 GB:M21277 GB:M55989 GB:M96795
PID:146181; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11

```

gene	/product="aerobic glycerol-3-phosphate dehydrogenase"	
	/protein_id="AAF96558.1"	
CDS	/db_xref="GI:9658073"	
	/translation="MNARSRFTQDEVLILIVGGGNGAGIAADATGRGLKVLGYLDA KDFASATSSASGLVHGLKYLEYBFLVSEALAEVLLKAPHLVTPMRFPLPHR PFLRPAMWIRAGLFLYDNLGKRTSLPASHKVNLIKAGSVTKPEMOIGFEYSDCWDDAR LVILNMAQOAGAEVLNCTVKEAKRMGDLHVTLDEOQOQFERRSHALVNAAGP WVQFLNHAHVSPYGIIRLIQSGSHIIVPRLHDEPOAYILONEDKRIYFVPIYLLDDYS MIGTIDVEYKDPKVAITDAERDYLISIVNKHMRILARSDDIAEFSGVRLPDCDES NSPOAITRDYTLSDQDAEAPLISIFGGKUTTYRKLGEAAWKHLAPFFPKMKAPWIA DAPLGGENFDYALKQLVAAFPFITEVIERWLRSYGSRTQLLAGVTGIEDLGLA FSGSLYQREIDYLCERKEFARHAQDIFWRRSLKGLNHDTSVVVEVESYLIQKFHAEQPL KATV"	
gene	complement (7648..9258)	
	/gene="VCA0658"	
CDS	complement (7648..9258)	
	/gene="VCA0658"	
gene	/note="similar to GB:X61367 PID:48195; identified by sequence similarity; putative"	
	/codon_start=1	
gene	/transl_table=11	
	/product="methyl-accepting chemotaxis protein"	
CDS	/protein_id="AAF96559.1"	
	/db_xref="GI:9658074"	
gene	/translation="MYTFALIKITLQQAEMSAVTPSAQOEVLVGDHDLVSTDLKG VITYCNDTPFCFIAGQDELKGNENIVRHASMPKAAFADMMHVLKQGHAMRGIVKNR TKSGPYWADVPIYQOQLTGQSVRVVRAERKWEIATYKAYQALLAAKAGKKIQ FKLHSLRYALLGALSPALAHQPAEQWOLASLRITILGRMDSARPLBLANQLH KQWNEYDSISRLYSGADAFSVADYHLKMASARITILGRMDSARPLBLANQLH TQVEHQALAAQNIQAVQATDAVEASAERVSHTSAHQILDOVDHCAETKHSI NVTHNQLRATAEBSAALTTLKLSDAQOQVQLMTEIGGIAEOTNLILNAALEAAR AGEGRGVADVDEBSARTORATQIOTSIDTLMSTIEAWRGDITASRDQTEOCA QDANTLLOQLQVECVMSDMLRVIGHVASAAHQRELTCENVQNHISATVATQNSAA THTVQLAMVSGKVAEFGALSKQPAQK"	
	complement (9444..10067)	
gene	/gene="VCA0659"	
	complement (9444..10067)	
CDS	/gene="VCA0659"	
	/note="similar to PID:2062640; identified by sequence similarity; putative"	
gene	/codon_start=1	
	/transl_table=11	
CDS	/product="protein F-related protein"	
	/protein_id="AAF96560.1"	
gene	/db_xref="GI:9658075"	
	/translation="MFEMKFWLVCVTFILLTCGSLVTDPMFNLLDVAPKDPDVRV PEWGYAPKNTSMVPMNTQPKRSYEDRLQYLLNSGLQHEVVPDGYEMIKLNNTVRF ETGSAKVSMAKQWLDTVARFLATESGIDIVLEGTNDTNGSEKLNKLARRANAVKA ALVQSRVAQNAIYTRGFGENVPACTNSTKNGRACNRRVEIRFILASN"	
gene	complement (10068..10190)	
	/gene="VCA0660"	
CDS	complement (10068..10190)	
	/gene="VCA0660"	
gene	/note="identified by Glimmer2; putative"	
	/codon_start=1	
CDS	/transl_table=11	
	/product="hypothetical protein"	
gene	/protein_id="AAF96561.1"	
	/db_xref="GI:9658076"	
CDS	/translation="MNGEHSNHLCDGKQPKICKLLMLTDKVFNNPINKYKNN"	

Search completed: October 30, 2004, 15:50:51

Job time : 6350 secs

ORIGIN	
Query Match 15.2%; Score 231.8; DB 1; Length 10264;	
Best Local Similarity 55.4%; Pred.No.4.9e-51;	
Matches 497; Conservative 0; Mismatches 387; Indels 13; Gaps 2;	
Qy	3 CATGGCATCTGCGCGTTCGCGTTCCTGACAGTGTGGTGGTTTCACGCAACCAAGCG 62
Db	1071 CTTCCGGAATGCCCTTTTGTATTGCTGCTGATTGTTGGCATTTTCTGCGCTCGGCAA 1012
Qy	63 TTTCCGCGCAATAGTTCCTGGCGCGCGATGATGCGATGCGATGCGTGTCCCGAGCT 122
Db	1011 GTTTGGCGGTAACCCCTTTTGGCGCTGCG-CTGGGTATGCTGATGGTGCATCCCGATC 953

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 30, 2004, 12:39:21 ; Search time 726 Seconds
(without alignments)
11041.138 Million cell updates/sec

Title: US-09-604-231-1

Perfect score: 1527

Sequence: 1 ctcattggcattcgcgcgtt.....gttgaaaccttgagtgttcg 1527

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4134886 segs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_23Sep04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	ID	Description
1	1527	100.0	1527	5	Aaf31528 C.glutami
2	1527	100.0	1527	6	Abs65346 DNA encod
3	1515	99.2	349980	5	Aah68533 C glutami
4	1492	97.7	1983	5	Aah67869 C glutami
5	1412.6	92.5	5969	4	Aaf32543 Brevibact
6	1107.4	72.5	1109	5	Aaf31529 C.glutami
7	218.8	14.3	110000	10	Continuation (6 of
8	190.6	12.5	2913	4	Aah54100 S. epider
9	177.8	11.6	975	6	Abk73538 Bacillus
10	174.6	11.4	861	8	Acf74550 Staphyloc
11	171.6	11.2	1491	10	Adh84532 Enterococ
12	151.2	9.9	5840	2	Aax12968 Enterococ
13	151.2	9.9	5840	6	Abss98763 Enterococ
14	143.6	9.4	1386	11	ABD000428
15	143.6	9.4	1452	11	ABD00072 Klebsiell
16	137.4	9.0	732	6	ABN91814
17	137.2	9.0	1380	6	Abk73394 Bacillus
18	135.4	8.9	3037	6	Abq70786 Listeria
19	134.6	8.8	110000	6	Continuation (3 of
20	134.6	8.8	110000	6	Abq69245 Listeria
21	132.2	8.7	110000	6	Abq03041 Listeria

22	126.6	8.3	1965	10	ABX07668
23	126.6	8.3	2127	8	ABZ42306
24	126.6	8.3	3895	2	AAV52334
25	126.6	8.3	110000	10	ABS56454_16
26	121.4	8.0	2028	6	ABN66285
27	121.4	8.0	110000	6	ABN71527_01
28	114.6	7.5	951	10	ADH84192
29	113.4	7.4	1884	4	AA555732
30	111	7.3	1543	10	ADD13201
31	109.6	7.2	3615	8	ABZ77360
32	109.6	7.2	3615	10	AAU51848
33	109.4	7.2	1905	10	ADF02562
34	108	7.1	3615	2	AAQ55752
35	107.8	7.1	1185	8	ABZ42298
36	107.8	7.1	9769	2	AAV52163
37	107.8	7.1	110000	10	ABS56454_15
38	105.6	6.9	1014	6	ABK73531
39	103.6	6.8	1917	6	ABN67846
40	103.6	6.8	110000	6	ABN71527_16
41	99	6.5	1098	6	ABK73443
42	98	6.4	489	6	ABK73566
43	97.8	6.4	465	2	AAI47502
44	97.2	6.4	1971	11	ACH95246
45	93.6	6.1	1854	6	ABQ69909

ALIGNMENTS

RESULT 1

AAF31528
ID AAF31528 standard; DNA; 1527 BP.
AC AAF31528;
XX
DT 09-APR-2001 (first entry)
XX
DE C.glutamicum phosphoenolpyruvate DNA #1.
XX
KW Phosphoenolpyruvate; sugar phosphotransferase system; PTS; ds.
XX
OS Corynebacterium glutamicum.
XX
PN WO200102583-A2.
XX
PD 11-JAN-2001.
XX
PF 27-JUN-2000; 2000WO-IB000973.
XX
PR 01-JUL-1999; 99US-0142691P.
PR 23-AUG-1999; 99US-0150310P.
PR 03-SEP-1999; 99DE-01042095.
PR 03-SEP-1999; 99DE-01042097.
(BADI) BASF AG.
Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;
WPI; 2001-080989/09.
XX
Corynebacterium glutamicum nucleic acids encoding phosphoenolpyruvate:
sugar phosphotransferase system proteins or their portions, useful for
typing or identifying C. glutamicum or related bacteria, and as markers
for transformation.
Claim 3; Page 98-101; 144pp; English.
XX
The present invention relates to Corynebacterium glutamicum
phosphoenolpyruvate: sugar phosphotransferase system (PTS) proteins. The
PTS nucleic acids and proteins are useful in the identification of
microorganisms which can be used to produce fine chemicals, for
modulating fine chemical production in C. glutamicum or related bacteria,
the typing or identification of C. glutamicum or related bacteria, as

Abx07668 S. pneumo
Abz42306 Streptoco
Aav52334 Streptoco
Continuation (17 o
Abn66285 Streptoco
Continuation (2 of
Adh84192 Streptoco
Aas55732 Streptoco
Add13201 C. glutam
Abz77360 Nucleotid
Aal51848 Escherich
Adf02562 Bacterial
Aaq55752 Escherich
Abz42298 Streptoco
Aav52163 Streptoco
Continuation (16 o
Abk73531 Bacillus
Abn67846 Streptoco
Continuation (17 o
Abk73443 Bacillus
Abk73566 Bacillus
Aat47502 Partial P
Ach95246 Klebsiell
Abq69909 Listeria

CC	reference points for mapping C. glutamicum genome, and as markers for transformation									
XX	Sequence	1527 BP;	304 A;	392 C;	430 G;	401 T;	0 U;	0 Other;		
QY	Query Match	100.0%; Score 1527; DB 5; Length 1527;								
XX	Best Local Similarity	100.0%; Pred. No. 0;								
QQ	Matches 1527; Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;		
QY	1	CTCATGGCAATGCGCCGCTTCCGCTTCTGTCAGTGTTCGCAAGTGTGGTTTACCGCAACCAAG	60							
DB	1	CTCATGGCAATGCGCCGCTTCCGCTTCTGTCAGTGTTCGCAAGTGTGGTTTACCGCAACCAAG	60							
QY	61	CGTTTCGGCGCAATGAGTTCCTGGCGCGCGGATTCGATGCGATGTTTCCCGAG	120							
DB	61	CGTTTCGGCGCAATGAGTTCCTGGCGCGCGGATTCGATGCGATGTTTCCCGAG	120							
QY	121	CTTGTGAACGGCTACGACGTGGCGCCACCATGCTCGGGCGGAAATGCCAATGTGTC	180							
DB	121	CTTGTGAACGGCTACGACGTGGCGCCACCATGCTCGGGCGGAAATGCCAATGTGTC	180							
QY	181	CCTGTTTGGTTAGATGTTGCCAAGCGGTACCGGCGCAACCGTTCCTGTGCTGCT	240							
DB	181	CCTGTTTGGTTAGATGTTGCCAAGCGGTACCGGCGCAACCGTTCCTGTGCTGCT	240							
QY	241	GGTTTCTTGGATTTCTGGCAACGATCGAAGTTCCTGACAGCGACTCAAGGGCACTGC	300							
DB	241	GGTTTCTTGGATTTCTGGCAACGATCGAAGTTCCTGACAGCGACTCAAGGGCACTGC	300							
QY	301	AGACTTCTGTGACTCCAGTCTGACGTTGCTGCTCACCAGGATTCCTTACATTCATCGC	360							
DB	301	AGACTTCTGTGACTCCAGTCTGACGTTGCTGCTCACCAGGATTCCTTACATTCATCGC	360							
QY	361	CATTGGCCAGCAATGCGTGGGTGGCGATGCTGTCGACACCGGCTACAGGGGACTTTA	420							
DB	361	CATTGGCCAGCAATGCGTGGGTGGCGATGCTGTCGACACCGGCTACAGGGGACTTTA	420							
QY	421	TCATTTTGGTGTCCAGTGGCGGCTGCTCTTCTGGTCTGCTCTACTCCCAATGCTCAT	480							
DB	421	TCATTTTGGTGTCCAGTGGCGGCTGCTCTTCTGGTCTGCTCTACTCCCAATGCTCAT	480							
QY	481	CACCTGTCTGCACCAAGTCTCCCGCAATGAGCTGGAGCTGTGTTAACAGGGTGGATC	540							
DB	481	CACCTGTCTGCACCAAGTCTCCCGCAATGAGCTGGAGCTGTGTTAACAGGGTGGATC	540							
QY	541	CTTTCATCTCGCAACGGCAATGAGCTTAATCGCCACGGGTGGCGCATGTTTGGCAT	600							
DB	541	CTTTCATCTCGCAACGGCAATGAGCTTAATCGCCACGGGTGGCGCATGTTTGGCAT	600							
QY	601	GTTCTTCTCGGCAAGAGTGAAGAGCTCAAGGCGCTTGCAGGTGCTTCAAGGTGCTCCGC	660							
DB	601	GTTCTTCTCGGCAAGAGTGAAGAGCTCAAGGCGCTTGCAGGTGCTTCAAGGTGCTCCGC	660							
QY	661	TGTTCTTGTGTTTACGGAGCCTGCGATCTTCGGTGTGAACCTTCGCTCGGCTGGCGGTT	720							
DB	661	TGTTCTTGTGTTTACGGAGCCTGCGATCTTCGGTGTGAACCTTCGCTCGGCTGGCGGTT	720							
QY	721	CTTTCATCGTATCGGTACCGCAGCTATCGGTGCGGCTTGTGATGCACTCTTTAATATCAA	780							
DB	721	CTTTCATCGTATCGGTACCGCAGCTATCGGTGCGGCTTGTGATGCACTCTTTAATATCAA	780							
QY	781	GGCAGTGTGCTTGGGCGCTGCAGGTTCTTGGGTGTTGTTTCTATTGATGCTCCAGATAT	840							
DB	781	GGCAGTGTGCTTGGGCGCTGCAGGTTCTTGGGTGTTGTTTCTATTGATGCTCCAGATAT	840							
QY	841	GGTCATGTTCTTGGTGTGCAATGTTTACCTTCTTCATCGCATTCGGCGCAGCGATTGC	900							
DB	841	GGTCATGTTCTTGGTGTGCAATGTTTACCTTCTTCATCGCATTCGGCGCAGCGATTGC	900							
QY	901	TTATGGCCTTTACTTGTTCGCGCAACGCGCATTTGATCCAGATGCAACCGCTGCTCC	960							
DB	901	TTATGGCCTTTACTTGTTCGCGCAACGCGCATTTGATCCAGATGCAACCGCTGCTCC	960							

RESULT 2

ABS65346

ID ABS65346 standard; DNA; 1527 BP.

XX ABS65346;

XX ABS65346;

XX 15-NOV-2002 (first entry)

XX DNA encoding C. glutamicum metabolic pathway (MP) protein #5.

XX Metabolic pathway protein; MP; cell metabolism; amino acid; vitamin;

XX cofactor; nucleotide; nucleoside; trehalose; fine chemical production;

XX organic acid; non-proteinogenic amino acid; purine base; carbohydrate;

XX pyrimidine base; lipid; unsaturated fatty acid; diols; polyketide;

XX aromatic compound; food industry; animal feed; cosmetic industry;

XX pharmaceutical industry; gene; ds.

XX Corynebacterium glutamicum ATCC 13032.

XX WO200251231-A1.

XX 04-JUL-2002.

XX 22-DEC-2000; 2000WO-BP013143.

XX 22-DEC-2000; 2000WO-BP013143.

XX (BADI) BASF AG.

XX Pompejus M, Kroeger B, Zelder O, Schroeder H;

XX WPI; 2002-643289/69.

DR P-PSDB; ABG80325.
XX New metabolic pathway genes of *Corynebacterium glutamicum* for producing
PT fine chemicals, e.g. lipids, (un)saturated fatty acids, vitamins,
PT cofactors or enzymes used in food, feed, cosmetics or pharmaceutical
PT industries.
XX
PS Claim 1; Page 108-111; 176pp; English.
XX
CC The present invention relates to the isolation of *Corynebacterium*
CC glutamicum metabolic pathway (MP) proteins, and the polynucleotide
CC sequences encoding them. The MP proteins are enzymes involved in the
CC metabolism of molecules important for the normal functioning of cells
CC (e.g. amino acids, vitamins, cofactors, nucleotides and nucleosides, or
CC trehalose). The polynucleotide sequences encoding the MP proteins are
CC useful for producing fine chemicals, particularly organic acids, non-
CC proteinogenic amino acids, purine and pyrimidine bases, nucleosides,
CC nucleotides, lipids, (un)saturated fatty acids, diols, carbohydrates,
CC aromatic compounds, vitamins, cofactors, polyketides and enzymes. The
CC fine chemicals are useful in the food, animal feed, cosmetic or
CC pharmaceutical industries. ABS65342-ABS65364 encode the *C. glutamicum* MP
CC proteins of the invention.
XX
SQ Sequence 1527 BP; 304 A; 392 C; 430 G; 401 T; 0 U; 0 Other;
Query Match 100.0%; Score 1527; DB 6; Length 1527;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1527; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTCATGGCATCTGCGCGTTCGGTCTTCGCGAGTGTGGTTCACCGCAACAAG 60
DB 1 CTCATGGCATCTGCGCGTTCGGTCTTCGCGAGTGTGGTTCACCGCAACAAG 60
QY 61 CGTTTCGGCGCAATCAGATTCTCGCGCGCGTATTGGTATGCGATGTTCCCGAG 120
DB 61 CGTTTCGGCGCAATCAGATTCTCGCGCGCGTATTGGTATGCGATGTTCCCGAG 120
QY 121 CTTGGTGAACGGCTACGACGTGCGCGCCACCATGGCTGCGGGCAAAATGCCATGTGGTC 180
DB 121 CTTGGTGAACGGCTACGACGTGCGCGCCACCATGGCTGCGGGCAAAATGCCATGTGGTC 180
QY 181 CTTGTTTGGTTTATGTTGCGCAACCGGTTACCAAGGACCGTCTTCCTGTGTGGT 240
DB 181 CTTGTTTGGTTTATGTTGCGCAACCGGTTACCAAGGACCGTCTTCCTGTGTGGT 240
QY 241 GGTTCCTTCGATTCTCGCAACGATCAGAGATTCTTCGCAACGACTCAAGGGCACTGC 300
DB 241 GGTTCCTTCGATTCTCGCAACGATCAGAGATTCTTCGCAACGACTCAAGGGCACTGC 300
QY 301 AGACTTCCTGATCACTCCAGTGTGACGTGTGCTGCTCACCGGATTCCTTACATTCATCGC 360
DB 301 AGACTTCCTGATCACTCCAGTGTGACGTGTGCTGCTCACCGGATTCCTTACATTCATCGC 360
QY 361 CATTCGCCAGCAATGCGTGGTGGCGATGTCTGGCACACGCTCTACAGGACTTTA 420
DB 361 CATTCGCCAGCAATGCGTGGTGGCGATGTCTGGCACACGCTCTACAGGACTTTA 420
QY 421 TGAATTCGGTGTGTCAGTCCGGGTCTGTCTTCCTCGGTCTGTCTACTCCCAATCGTCAT 480
DB 421 TGAATTCGGTGTGTCAGTCCGGGTCTGTCTTCCTCGGTCTGTCTACTCCCAATCGTCAT 480
QY 481 CACTGGTCTGCAACCAATCTCCCGCAATGAGCTGGAGCTGTTTAAACAGGGTGGATC 540
DB 481 CACTGGTCTGCAACCAATCTCCCGCAATGAGCTGGAGCTGTTTAAACAGGGTGGATC 540
QY 541 CTTTCATCTTCGCAACGGCATCTATGCTAATATCGGCCAGGGTGGCGCATGTTTGGCAGT 600
DB 541 CTTTCATCTTCGCAACGGCATCTATGCTAATATCGGCCAGGGTGGCGCATGTTTGGCAGT 600
QY 601 GTTCTTCCTCGCGCAAGAGTGAAGCTCAAGGCTTCAGGCTTCAGGCTTCCTCCCG 660
DB 601 GTTCTTCCTCGCGCAAGAGTGAAGCTCAAGGCTTCAGGCTTCAGGCTTCCTCCCG 660

QY 661 TGTTCCTTGGTATTACGAGCGCTCGATCTTCGGTGTGAACCTTCGCGCTCGCGTGGCGGTT 720
DB 661 TGTTCCTTGGTATTACGAGCGCTCGATCTTCGGTGTGAACCTTCGCGCTCGCGTGGCGGTT 720
QY 721 CTTTCATCGGTATCGGTACCGCAGCTATCGGTGCGGCTTTGATTGCACTCTTTAATATCAA 780
DB 721 CTTTCATCGGTATCGGTACCGCAGCTATCGGTGCGGCTTTGATTGCACTCTTTAATATCAA 780
QY 781 GGCAGTTCGGTTCGGCGCTGCGAGTTCTTGGGTGTGTTTCTTATGATGCTCCAGATAT 840
DB 781 GGCAGTTCGGTTCGGCGCTGCGAGTTCTTGGGTGTGTTTCTTATGATGCTCCAGATAT 840
QY 841 GGTTCATGTTCTTCGGTGTGCGAGTTCTTCTTCATCGCATTCGGCGCAGCATTCG 900
DB 841 GGTTCATGTTCTTCGGTGTGCGAGTTCTTCTTCATCGCATTCGGCGCAGCATTCG 900
QY 901 TTATGCGCTTTACTTGGTTCGGCGCAACCGCGCAGCATTCGATCCAGATGCAACCGCTGCTC 960
DB 901 TTATGCGCTTTACTTGGTTCGGCGCAACCGCGCAGCATTCGATCCAGATGCAACCGCTGCTC 960
QY 961 AGTGCCTGCGAGGAACGACCAAGCCGAGCAGNAGCACCCGCGAGATTTTCAACGATTC 1020
DB 961 AGTGCCTGCGAGGAACGACCAAGCCGAGCAGNAGCACCCGCGAGATTTTCAACGATTC 1020
QY 1021 CACCATCATCCAGGCACCTTTTGACCGGTGAAGCTATTGCACTCAGCAGCGTCAGCGATGC 1080
DB 1021 CACCATCATCCAGGCACCTTTTGACCGGTGAAGCTATTGCACTCAGCAGCGTCAGCGATGC 1080
QY 1081 CATGTTTGCAGCGGAAAGCTTGGCTCGGCGGTGCGCATCGTCCCAACCAAGGGGCGAGTT 1140
DB 1081 CATGTTTGCAGCGGAAAGCTTGGCTCGGCGGTGCGCATCGTCCCAACCAAGGGGCGAGTT 1140
QY 1141 AGTTTCTCCGCTGAGTGGAAAGATTGTGTGGCATTCCTCATCTGGCGCATGCTTTCGAGT 1200
DB 1141 AGTTTCTCCGCTGAGTGGAAAGATTGTGTGGCATTCCTCATCTGGCGCATGCTTTCGAGT 1200
QY 1201 TCGCAACCAAGGCTGAGGATGTTCCAAATGTGGATATCTTGTATGCATATGGTTTCGACAC 1260
DB 1201 TCGCAACCAAGGCTGAGGATGTTCCAAATGTGGATATCTTGTATGCATATGGTTTCGACAC 1260
QY 1261 AGTAAACCTCAACGGCAGCAGCTTTAAACCGCTGGAAGAAGCAGGCGATGAAGTCAAAGC 1320
DB 1261 AGTAAACCTCAACGGCAGCAGCTTTAAACCGCTGGAAGAAGCAGGCGATGAAGTCAAAGC 1320
QY 1321 AGGGGAGCTGCTGTGTAATTCGATATTCATGCGCATTAAGGCTGAGGTTATGAGGTAAC 1380
DB 1321 AGGGGAGCTGCTGTGTAATTCGATATTCATGCGCATTAAGGCTGAGGTTATGAGGTAAC 1380
QY 1381 CACGCGGATTTGTTTTCGAATTAACAAGAAACCGGACCTGTAAACACTTACCGTTTGGG 1440
DB 1381 CACGCGGATTTGTTTTCGAATTAACAAGAAACCGGACCTGTAAACACTTACCGTTTGGG 1440
QY 1441 CGAAATTTGAAGCGGAGCCAACTGCTCAACGTCGCAAGAAAGAAAGCGGTGCCAGCAAC 1500
DB 1441 CGAAATTTGAAGCGGAGCCAACTGCTCAACGTCGCAAGAAAGAAAGCGGTGCCAGCAAC 1500
QY 1501 ACCATAAGTTGAACCTTGAGTGTTCG 1527
DB 1501 ACCATAAGTTGAACCTTGAGTGTTCG 1527

RESULT 3

AAH68533/c
ID AAH68533 standard; DNA; 349980 BP.

XX AAH68533;

AC

XX

DT 26-SEP-2001 (first entry)

XX

DE C glutamicum coding sequence fragment SEQ ID NO: 7068.

XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;

KW organic acid synthesis; ds.


```
XX Corynebacterium glutamicum.
OS EP1108790-A2.
PN 20-JUN-2001.
XX 18-DEC-2000; 2000EP-00127688.
XX 16-DEC-1999; 99JP-00377484.
PR 07-APR-2000; 2000JP-00159162.
PR 03-AUG-2000; 2000JP-00280988.
XX (KYOW ) KYOWA HAKKO KOGYO KK.
PA
XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX WPI; 2001-376931/40.
XX Novel polynucleotides derived from Coryneform bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analyzing
PT expression profile or pattern of a gene and identifying homologous gene.
XX Disclosure; SEQ ID NO 7068; 246pp + Sequence Listing; English.
XX
XX The present invention provides a number of nucleotide and protein
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and analysing
CC the expression profile or expression pattern of a gene derived from
CC Coryneform bacterium, and identifying a homologue of a gene derived from
CC Coryneform bacterium. Coryneform bacteria are useful for producing amino
CC acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a nucleic acid, described
CC in the exemplification of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from the European Patent Office
XX
XX Sequence 349980 BP; 81250 A; 97718 C; 90621 G; 80391 T; 0 U; 0 Other;
SQ
Query Match 99.2%; Score 1515; DB 5; Length 349980;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1526; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1 CTCATGGCAATCTGCGCGGTCGCGTTCCTTCCAGTGTTCGTTGTTTCCCGCAACCAAG 111267
DB 111326 CTCATGGCAATCTGCGCGGTCGCGTTCCTTCCAGTGTTCGTTGTTTCCCGCAACCAAG 111267
QY 61 CGTTTCGCGGCAATGAGTTCCTGCGCGCGGCTATTGTTGATGGCGATGTTCCCGAG 120
DB 111266 CGTTTCGCGGCAATGAGTTCCTGCGCGCGGCTATTGTTGATGGCGATGTTCCCGAG 111208
QY 121 CTTGTGTAAACGGCTACGAGTGTGCGCGCCACCATGCTGCGGCGGAAATCCCAATGTGTC 180
DB 111207 CTTGTGTAAACGGCTACGAGTGTGCGCGCCACCATGCTGCGGCGGAAATCCCAATGTGTC 111148
QY 181 CCGTGTGTTGTTAGATGTTGCCAAGCCGTTACAGGSCACCGTCTTCCTGTGCTGCT 240
DB 111147 CCGTGTGTTGTTAGATGTTGCCAAGCCGTTTACAGGSCACCGTCTTCCTGTGCTGCT 111088
QY 241 GGTTCCTCTGATTCCTGCAACGATCGAGAAGTTCCTGCAACAGCGACTCAAGGGCACTGC 300
DB 111087 GGTTCCTCTGATTCCTGCAACGATCGAGAAGTTCCTGCAACAGCGACTCAAGGGCACTGC 111028
QY 301 AGACTTCCTGATCACTCCAGTGTGTCAGTGTGCTGCTACCGGATTCCTTACATTCATCCG 360
DB 111027 AGACTTCCTGATCACTCCAGTGTGTCAGTGTGCTGCTACCGGATTCCTTACATTCATCCG 110968
QY 361 CATTTGCCCGCAGCAATCGGCTGGGCGGATGCTGGCACACGGCTACAGGACTTTTA 420
DB 110967 CATTTGCCCGCAGCAATCGGCTGGGCGGATGCTGGCACACGGCTACAGGACTTTTA 110908
QY
421 TGATTTCCGTTGGTCCAGTCCGCGGTCGCTCTTCCGTTGGTGTCTACTACCAATCGTCAT 480
DB 110907 TGATTTCCGTTGGTCCAGTCCGCGGTCGCTCTTCCGTTGGTGTCTACTACCAATCGTCAT 110848
QY 481 CACTGGTCTGCACCAAGTCTTCCCGCAATTTGAGCTGGAGCTGTTTAAACCAAGGTTGATC 540
DB 110847 CACTGGTCTGCACCAAGTCTTCCCGCAATTTGAGCTGGAGCTGTTTAAACCAAGGTTGATC 110788
QY 541 CTTTCATCTTCCCAAGCGCATCTATGGCTAAATATGCCCCAGGTCGGCATGTTTGGCAGT 600
DB 110787 CTTTCATCTTCCCAAGCGCATCTATGGCTAAATATGCCCCAGGTCGGCATGTTTGGCAGT 110728
QY 601 GTTCTTCTCGCGAAGAGTGAAGAGCTCAAGGCGCTTCAGGTGCTTCAGGTGCTCCGC 660
DB 110727 GTTCTTCTCGCGAAGAGTGAAGAGCTCAAGGCGCTTCAGGTGCTTCAGGTGCTCCGC 110668
QY 661 TGTTCCTCGATATACGAGAGCTGCGATCTTCGTTGTGAACCTTCGCTTCGCTGCGCTGCTT 720
DB 110667 TGTTCCTCGATATACGAGAGCTGCGATCTTCGTTGTGAACCTTCGCTTCGCTGCGCTGCTT 110608
QY 721 CTTTCATCGTATCGGTACCGAGCTATCGGTGGCGCTTGTGATGCACTCTTTAATATCAA 780
DB 110607 CTTTCATCGTATCGGTACCGAGCTATCGGTGGCGCTTGTGATGCACTCTTTAATATCAA 110548
QY 781 GGCAGTTCGTTGGCGCTGCGAGTTCCTTGGGTGTTTCTTATGATGCTCCAGATAT 840
DB 110547 GGCAGTTCGTTGGCGCTGCGAGTTCCTTGGGTGTTTCTTATGATGCTCCAGATAT 110488
QY 841 GGTCAATGTTCTTGGTGTGTCAGTGTTCCTTCTCATCGCATTCGCGGAGCGGATTCG 900
DB 110487 GGTCAATGTTCTTGGTGTGTCAGTGTTCCTTCTCATCGCATTCGCGGAGCGGATTCG 110428
QY 901 TTATGGCTTCTTACTTGGTTCGCGCAACGGCAGCATTCATCCAGATGCAACCGCTGCTCC 960
DB 110427 TTATGGCTTCTTACTTGGTTCGCGCAACGGCAGCATTCATCCAGATGCAACCGCTGCTCC 110368
QY 961 AGTGCTCGAGGAACGACAAAGCCGAGCAGAGACCCGAGAAATTTTCAAGATTC 1020
DB 110367 AGTGCTCGAGGAACGACAAAGCCGAGCAGAGACCCGAGAAATTTTCAAGATTC 110308
QY 1021 CACCATCATCCAGGACACCTTTGACCGGTGAGCTATTGCACTGAGCAGCGTCAGCGATGC 1080
DB 110307 CACCATCATCCAGGACACCTTTGACCGGTGAGCTATTGCACTGAGCAGCGTCAGCGATGC 110248
QY 1081 CATGTTTCCAGCGGAAAGCTTGGCTCGGGCGTTCGCAATCGTCCCAACCAAGGGGAGTT 1140
DB 110247 CATGTTTCCAGCGGAAAGCTTGGCTCGGGCGTTCGCAATCGTCCCAACCAAGGGGAGTT 110188
QY 1141 AGTTTCTCCGTTGAGTGAAGAGTTCGTTGGTGCATTCCTCCATCTGGCCATGCTTTCGAGT 1200
DB 110187 AGTTTCTCCGTTGAGTGAAGAGTTCGTTGGTGCATTCCTCCATCTGGCCATGCTTTCGAGT 110128
QY 1201 TCGCACCAAGCTGAGGATGTTTCCATGTGATATCTTGTGATGTCATGTTGTTCCACAC 1260
DB 110127 TCGCACCAAGCTGAGGATGTTTCCATGTGATATCTTGTGATGTCATGTTGTTCCACAC 110068
QY 1261 AGTAAACCTCAACCGGACGACCTTTAAACCCGCTGAAAGAGCAGGCGGATGAAGTCAAGC 1320
DB 110067 AGTAAACCTCAACCGGACGACCTTTAAACCCGCTGAAAGAGCAGGCGGATGAAGTCAAGC 110008
QY 1321 AGGGAGCTGCTGTGTGAATTCGATATTCGATTCATTAAGGCTGCAAGTTATGAGTAAAC 1380
DB 110007 AGGGAGCTGCTGTGTGAATTCGATATTCGATTCATTAAGGCTGCAAGTTATGAGTAAAC 109948
QY 1381 CACGCGGATGTTGTTTCCAAATTAAGAAACCGGACCTGTAAACACATTCAGCTTTGG 1440
DB 109947 CACGCGGATGTTGTTTCCAAATTAAGAAACCGGACCTGTAAACACATTCAGCTTTGG 109888
QY 1441 CGAAATTTGAAGCGGAGCGCAACCTGCTCAACGTCGCAAAAGAAAGACCGGTCCAGCAAC 1500
DB 109887 CGAAATTTGAAGCGGAGCGCAACCTGCTCAACGTCGCAAAAGAAAGACCGGTCCAGCAAC 109828
QY 1501 ACCATAAGTTGAACACCTTGAGTGTTCG 1527
```

Db 109827 ACCATAAGTTGAAACCTTGAGTGTTCG 109801
|||||
RESULT 4
AAH67869
ID AAH67869 standard; DNA; 1983 BP.
XX AC
XX AAH67869;
XX
XX 26-SEP-2001 (first entry)
XX
XX C glutamicum coding sequence fragment SEQ ID NO: 2904.
XX
XX Corynebacterium; amino acid synthesis; vitamin; saccharide;
KW organic acid synthesis; ds.
XX
XX Corynebacterium glutamicum.
XX
XX BP1108790-A2.
XX
XX 20-JUN-2001.
XX
XX 18-DEC-2000; 2000BP-00127688.
XX
XX 16-DEC-1999; 99JP-00377484.
PR 07-APR-2000; 2000JP-00159162.
PR 03-AUG-2000; 2000JP-00280988.
XX
XX (KYOW) KYOWA HAKKO KOGYO KK.
XX
XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX
XX WPI; 2001-376931/40.
DR P-PSDB; AAG92650.
XX
XX Novel polynucleotides derived from Coryneform bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analyzing
PT expression profile or pattern of a gene and identifying homologous gene.
XX
XX Claim 8; SEQ ID NO 2904; 246pp + Sequence Listing; English.
XX
XX The present invention provides a number of nucleotide and protein
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and analysing
CC the expression profile or expression pattern of a gene derived from
CC Coryneform bacterium, and identifying a homologue of a gene derived from
CC Coryneform bacterium. Coryneform bacteria are useful for producing amino
CC acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a nucleic acid described
CC in the exemplification of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from the European Patent Office
XX
SQ Sequence 1983 BP; 410 A; 508 C; 558 G; 507 T; 0 U; 0 Other;
Query Match 97.7%; Score 1492; DB 5; Length 1983;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1503; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
1 CTCATGGCATCTGCGCGTTCGGTTCCTTCCAGTGTGGTGGTTTCACCGCAACCAAG 60
481 CTCATGGCATCTGCGCGTTCGGTTCCTTCCAGTGTGGTGGTTTCACCGCAACCAAG 540
61 CGTTTCGGCGGCAATGAGTTCCTGGCGCGCGGTATGTGATGGCGATGGTTCCTCCGAG 120
541 CGTTTCGGCGGCAATGAGTTCCTGGCGCGCGGTATGTGATGGCGATGGTTCCTCCGAG 599
121 CTTTGGTGAAGGCTAGAGTTCGGCGCGCGGTATGGTTCGGCGGCAATGCCAATGGTTC 180
600 CTTTGGTGAAGGCTAGAGTTCGGCGCGCGGTATGGTTCGGCGGCAATGCCAATGGTTC 659

QY 181 CCTGTTTGGTTTGAATGTTTCCCAAGCCGTTTACCAGGGGCAACCGTGTCTCTCTGTGCTGGT 240
Db 660 CCTGTTTGGTTTGAATGTTTCCCAAGCCGTTTACCAGGGGCAACCGTGTCTCTCTGTGCTGGT 719
QY 241 GGTTCCTTGGATTCTTGGCAACGATCGAGAAGTTCCTTGCAACAAGCGACTCAAGGGCACTGC 300
Db 720 GGTTCCTTGGATTCTTGGCAACGATCGAGAAGTTCCTTGCAACAAGCGACTCAAGGGCACTGC 779
QY 301 AGACTTCTTGATCATTCCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
Db 780 AGACTTCTTGATCATTCCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 839
QY 361 CATTTGCCCCAGCAATCGCTGGGTGGCGGATGCTGGGACACGGTCTACAGGACCTTTA 420
Db 840 CATTTGCCCCAGCAATCGCTGGGTGGCGGATGCTGGGACACGGTCTACAGGACCTTTA 899
QY 421 TGATTTCCGTTGCTCAGTCCGCGGTCTGCTCTTCCGGTCTTGGTCTTACTCAACAATCGTCAT 480
Db 900 TGATTTCCGTTGCTCAGTCCGCGGTCTGCTCTTCCGGTCTTGGTCTTACTCAACAATCGTCAT 959
QY 481 CACTGCTCTGCACCACTCTTCCCGCAATTGAGCTGGAGCTGTTTAAACCAAGGTGATC 540
Db 960 CACTGCTCTGCACCACTCTTCCCGCAATTGAGCTGGAGCTGTTTAAACCAAGGTGATC 1019
QY 541 CTTTCATCTTCGCAACGGCATCTATGCTAATATCGCCAGGGTGGCGCATGTTTGGCAGT 600
Db 1020 CTTTCATCTTCGCAACGGCATCTATGCTAATATCGCCAGGGTGGCGCATGTTTGGCAGT 1079
QY 601 GTTCTTCTGGCGAAGAGTGAAGGCTCAAGGSCCTTGAGGCTTTCAGGCTGCTCCCG 660
Db 1080 GTTCTTCTGGCGAAGAGTGAAGGCTCAAGGSCCTTGAGGCTTTCAGGCTGCTCCCG 1139
QY 661 TGTTCCTTGGTATTACCGAGCTGCGATCTTCCGTTGGAACCTTCGGCTGCGCTGGCGGTT 720
Db 1140 TGTTCCTTGGTATTACCGAGCTGCGATCTTCCGTTGGAACCTTCGGCTGCGCTGGCGGTT 1199
QY 721 CTTTCATCGTATCGGTACCGAGCTATCGGTGGCGCTTTCGATGTCATCTTTAATATCAA 780
Db 1200 CTTTCATCGTATCGGTACCGAGCTATCGGTGGCGCTTTCGATGTCATCTTTAATATCAA 1259
QY 781 GGCAGTTCGCTTGGCGCTGCGAGTTCCTTGGGTGTTGTTTCTATGATGCTCCAGATAT 840
Db 1260 GGCAGTTCGCTTGGCGCTGCGAGTTCCTTGGGTGTTGTTTCTATGATGCTCCAGATAT 1319
QY 841 GGTTCAGTTCTTGGTGTGCGAGTGTTCCTTTCATCGCATTCGGCGCGAGGATTC 900
Db 1320 GGTTCAGTTCTTGGTGTGCGAGTGTTCCTTTCATCGCATTCGGCGCGAGGATTC 1379
QY 901 TTATGGCTTTTACTTGGTTCGCGCAACGGCAGCATTTGATCCAGATGCAACCGCTGCTCC 960
Db 1380 TTATGGCTTTTACTTGGTTCGCGCAACGGCAGCATTTGATCCAGATGCAACCGCTGCTCC 1439
QY 961 AGTGCTGCAGGAACGACCAAGCCGAGCAGAGCAGACCCGAGATTTTCAACGATTC 1020
Db 1440 AGTGCTGCAGGAACGACCAAGCCGAGCAGAGCAGACCCGAGATTTTCAACGATTC 1499
QY 1021 CACCATCATCCAGGACCTTTGACCGGTGAAGCTATGCACTGAGCAGCGGTGAGCGATGC 1080
Db 1500 CACCATCATCCAGGACCTTTGACCGGTGAAGCTATGCACTGAGCAGCGGTGAGCGATGC 1559
QY 1081 CATGTTTGGCAGGGAAGCTTTGGCTCGGGCGTTCGATCCGATCTGGCCATGCTTTCGCACT 1140
Db 1560 CATGTTTGGCAGGGAAGCTTTGGCTCGGGCGTTCGATCCGATCTGGCCATGCTTTCGCACT 1619
QY 1141 AGTTTCTCCGTTGAGTGAAGATTTGTTGGTGCATTTCCCATCTGGCCATGCTTTCGCACT 1200
Db 1620 AGTTTCTCCGTTGAGTGAAGATTTGTTGGTGCATTTCCCATCTGGCCATGCTTTCGCACT 1679
QY 1201 TCGCACCAAGGCTGAGGATGTTTCAATGTGATATCTTGTGCAATGTTGTCATGTTTGCACAC 1260
Db 1680 TCGCACCAAGGCTGAGGATGTTTCAATGTGATATCTTGTGCAATGTTTGCATGTTTGCACAC 1739

QY 1261 AGTAACTCAACGCGACGCTTTAACCCTGGAAGAGCAGCGGATGAAGTCAAGC 1320
DB |||||
QY 1740 AGTAACTCAACGCGACGCTTTAACCCTGGAAGAGCAGCGGATGAAGTCAAGC 1799
DB |||||
QY 1321 AGGGAGCTGCTGTGTAATTCGATATTGATGCCATTAAAGCTGCAGGTTATGAGTTAAC 1380
DB |||||
QY 1800 AGGGAGCTGCTGTGTAATTCGATATTGATGCCATTAAAGCTGCAGGTTATGAGTTAAC 1859
DB |||||
QY 1381 CACGCCGATGTTGTTTCGAATTCAAGAAACCGGACCTGTAAACACTTACGTTTGGG 1440
DB |||||
QY 1860 CACGCCGATGTTGTTTCGAATTCAAGAAACCGGACCTGTAAACACTTACGTTTGGG 1919
DB |||||
QY 1441 CGAAATTCAGCGGAGCGCAACCTCTCAAGTTCGCAAGAAAGAGCGGTGCCAGCAAC 1500
DB |||||
QY 1920 CGAAATTCAGCGGAGCGCAACCTCTCAAGTTCGCAAGAAAGAGCGGTGCCAGCAAC 1979
DB |||||
QY 1501 ACCA 1504
DB |||||
QY 1980 ACCA 1983
DB |||||
RESULT 5
AAF32543
ID AAF32543 standard; DNA; 5969 BP.
XX
AC AAF32543;
XX
DT 11-SEP-2003 (revised)
DT 20-APR-2001 (first entry)
XX
DE Brevibacterium lactofermentum sucrose PTS enzyme II DNA SEQ ID NO:1.
XX
KW Brevibacterium lactofermentum; sucrose PTS enzyme II; sucrose;
KW phosphoenolpyruvate:carboxylate phosphotransferase system; glucose;
KW coryneform bacterium; phosphoenolpyruvate-sugar transport system; ds.
XX
OS Corynebacterium glutamicum.
XX
FH Key Location/Qualifiers
FT CDS 3779..5764
FT /*tag= a
FT /product= "sucrose PTS enzyme II"
XX
XX WO200102584-A1.
XX
PD 11-JAN-2001.
XX
XX 30-JUN-2000; 2000WO-JP004348.
XX
XX 02-JUL-1999; 99JP-00189512.
XX
XX (AJUN) AJINOMOTO CO INC.
XX
PI Izui M, Sugimoto M, Nakamatsu T, Kurahashi O;
XX
XX WPI: 2001-138150/14.
DR P-PDB; AAB69080.
XX
XX Gene encoding sucrose phosphoenolpyruvate-sugar transport system enzyme
PT II obtained by cassette ligation-mediated amplification of downstream
PT domain of coryneform bacterium sucrose gene, with sucrose-binding
PT activity.
XX
PS Claim 3; Page 22-29; 45pp; Japanese.
XX
XX The present sequence encodes the Brevibacterium lactofermentum sucrose
CC PTS (phosphoenolpyruvate:carboxylate phosphotransferase system or
CC phosphoenolpyruvate-sugar transport system) enzyme II, which has sucrose-
CC binding activity. A coryneform bacteria produced with the sucrose PTS
CC enzyme II gene can have more efficient sugar uptake, and improved amino-
CC acid and nucleic acid productivity. The sucrose PTS gene and it's
CC disrupted gene, such as one without the sucrose PTS function, can be used
CC to produce new breeds of coryneform bacterial strains to uptake sugar

CC more efficiently e.g. glucose only or and sucrose, and can have improved
CC amino-acid and nucleic acid productivity. (Updated on 11-SEP-2003 to
CC standardise OS field)
XX
SQ Sequence 5969 BP; 1388 A; 1575 C; 1578 G; 1428 T; 0 U; 0 Other;
Query Match 92.5%; Score 1412.6; DB 4; Length 5969;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 1462; Conservative 0; Mismatches 64; Indels 1; Gaps 1;
QY 1 CTCATGGCATCTGGCCGCTTCGGCTTCCTGCGGCGCGGATTCGATGGGATGGTTCCTCCGAG 60
DB CTGATGGCATCTGGCCGCTTCGGCTTCCTGCGGCGCGGATTCGATGGGATGGTTCCTCCGAG 4318
QY 61 CGTTTCGCGCAATGAGTTCCTGCGGCGCGGATTCGATGGGATGGTTCCTCCGAG 120
DB CGTTTCGCGCAATGAGTTCCTGCGGCGCGGATTCGATGGGATGGTTCCTCCGAG 4377
QY 121 CTTCGTGAACGGCTACGACGTCGGCGCCACCATGCTCGCGGCGGAAATGCCAATGTGTC 180
DB CTTCGTGAACGGCTACGACGTCGGCGCCACCATGCTCGCGGCGGAAATGCCAATGTGTC 4437
QY 181 CCTGTTTGGTTAGATGTTGCCAAGCCGTTTACAGGGCACCGTTCCTGCTGCTGCT 240
DB CCTGTTTGGTTAGATGTTGCCAAGCCGTTTACAGGGCACCGTTCCTGCTGCTGCTGCT 4497
QY 241 GGTTCCTTGGATTCGCGCAACGATCGAAGATTCCTGCAAGGACCTCAAGGGCACTGC 300
DB GGTTCCTTGGATTCGCGCAACGATCGAAGATTCCTGCAAGGACCTCAAGGGCACTGC 4557
QY 301 AGACTTCCTGATCACTCCAGTCTGACGTTGCTGCTCACCAGGATTCCTTACATTCATCGC 360
DB AGACTTCCTGATCACTCCAGTCTGACGTTGCTGCTCACCAGGATTCCTTACATTCATTCG 4617
QY 361 CATTCGCCAGCAATGCGCTGGTGGGATGCTGTCGACACCGCTTACAGGGACTTTA 420
DB CATTCGCCAGCAATGCGCTGGTGGGATGCTGTCGACACCGCTTACAGGGACTTTA 4677
QY 421 TGATTTTCGTTGCTCCAGTCGGCGGCTGCTGCTTTCGGTCTGCTACTCACCATTCATCAT 480
DB TGATTTTCGTTGCTCCAGTCGGCGGCTGCTGCTTTCGGTCTGCTACTCACCATTCATTCAT 4737
QY 481 CACTGCTCTGCACCACTCTTCCCGCAATTCAGTTCGAGCTGTTTCAACACAGGGTGGATC 540
DB CACTGCTCTGCACCACTCTTCCCGCAATTCAGTTCGAGCTGTTTCAACACAGGGTGGATC 4797
QY 541 CTTTCATCTTCGCAACCGCATCTATGGCTAATATCCGCCAGGCTCGGATGTTTGGCACT 600
DB CTTTCATCTTCGCAACCGCATCTATGGCTAATATCCGCCAGGCTCGGATGTTTGGCACT 4857
QY 601 GTTCTTCCTGGCGAAGAGTGAAGCTCAAGGGCTTCGAGGTGCTTCAGGTGCTCTCCGC 660
DB GTTCTTCCTGGCGAAGAGTGAAGCTCAAGGGCTTCGAGGTGCTTCAGGTGCTCTCCGC 4917
QY 661 TGTTCCTGTTATTCAGGACCTTCGATCTTCGGTGTGAACCTTCGCTCGCTGCGCTGCT 720
DB TGTTCCTGTTATTCAGGACCTTCGATCTTCGGTGTGAACCTTCGCTCGCTGCGCTGCT 4977
QY 721 CTTTCATCTTCGTTACCGGATTCGTCGGCGCTTCGAGGTGCTTCAGGTGCTCTTAAATCAA 780
DB CTTTCATCTTCGTTACCGGATTCGTCGGCGCTTCGAGGTGCTTCAGGTGCTCTTAAATCAA 5037
QY 781 GGCAGTTGCGTTGGGCGCTGCGAGGTTTCTTGGGTGTTTCTTATGATGCTCCAGATAT 840
DB GGCAGTTGCGTTGGGCGCTGCGAGGTTTCTTGGGTGTTTCTTATGATGCTCCAGATAT 5097
QY 841 GGTTCATGTTTGGTGTGCGATGTTTACCTTCTTCATCGCATTCGGCGGAGCGATTGC 900
DB GGTTCATGTTTGGTGTGCGATGTTTACCTTCTTCATCGCATTCGGCGGAGCGATTGC 5157
QY 901 TTATGCGCTTACTTGGTTTCGGCGCAACGCGAGCATTTGATCCAGATGCAACCGCTGCTCC 960
DB TTATGCGCTTACTTGGTTTCGGCGCAACGCGAGCATTTGATCCAGATGCAACCGCTGCTCC 5217

Db 82988 GTTACTAAGAGATAATTATTGTAAGTAAATCAAAA 83023

RESULT 8
AAH54100/c
ID AAH54100 standard; DNA; 2913 BP.
XX AC AAH54100;
XX DT 03-SEP-2001 (first entry)
XX DE S. epidermidis genomic polynucleotide sequence SEQ ID NO:3464.
XX KW Staphylococcus epidermidis SRI strain; infection; diagnosis; vaccination;
XX OS endocarditis; ds.
XX PN Staphylococcus epidermidis.
XX PD WO200134809-A2.
XX PF 17-MAY-2001.
XX PR 09-NOV-2000; 2000WO-US030782.
XX PR 09-NOV-1999; 99US-0164258P.
XX PA (GLAX) GLAXO GROUP LTD.
XX PI Kimmerly WJ;
XX PI WPI; 2001-316495/33.
XX PT Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
XX PS useful for vaccinating against infections, e.g. endocarditis.
XX PS Claim 8; Page 1030-1031; 2188pp; English.
XX CC AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis. (I)
CC and (II) can have antibacterial activity and therefore can be used in
CC vaccination. The nucleic acids (I) may be used to produce the S.
CC epidermidis polypeptides (II) via the production of vectors containing
CC them which are used to produce hosts cells which express the
CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC used to vaccinate subjects and to raise antibodies against the bacteria.
CC The polypeptides may also be used to assay for other inhibitors of their
CC activity and therefore identify compounds that may be used for the
CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
CC AAH5090 represent specifically claimed S. epidermidis genomic DNA
CC polynucleotide sequences from the present invention. AAH55091 to AAH55098
CC represent oligonucleotide sequences and primers which are used in the
CC exemplification of the present invention. N.B. The present invention
CC specifically claims all the polynucleotide sequences given in the
CC sequence listing of the present specification, however the sequence
CC listing only goes up to SEQ ID NO:4454 so even though sequences are given
CC in the disclosure for SEQ ID NO:4465 to 4472, no sequences are present
CC for SEQ ID NO:4455 to 4464
XX
XX SQ Sequence 2913 BP; 1072 A; 480 C; 419 G; 942 T; 0 U; 0 Other;
Query Match 12.5%; Score 190.6; DB 4; Length 2913;
Best Local Similarity 55.0%; Pred. No. 1.7e-45;
Matches 468; Conservative 0; Mismatches 364; Indels 19; Gaps 4;
XX 5 TGGCATCTCGCGGTTCGGTCTTCCAGTGTGGTTCACCGCAACCAACGCTT 64
Db 2690 TTGCAATGCACCTTTATTTACCAATACTTATTGGTTTATGTGACGCTAAGCGAT 2631
XX 65 TCGGGCGCAATGAGTTCCTCGGCGCGCGTATTTGGTATGGCGATGTGTTCGCGAGCTTG 124
XX 2630 TTGGTGGTAATCTCTATTATTAGTGGAGC-TCTAGGTATGATGACTTGTTCATCTGGAATG 2572
XX 125 GTGAACGGGTACGAGGTGCGCCGCCACCATGGTGGCGGGCGAA---ATGCCAATGTGGTCC 181

Db 2571 ATGAGTGCAATGATTTTCCAAAAGCTTTAGAGAAGAAAGCTATTCACACTGGAT 2512
QY 182 CTGTTTGGTTTAGATGTTCGCCAAGCGGTTACAGGGCACCGTGTCTTCTGTGCTGTG 241
Db 2511 GTCTTTGGGCTACATATTAATGAAGTAGGTTATCAGGGACAAGTATTACCTATGCTCGTA 2452
QY 242 GTTTCTTGGATCTCGCAAGCATCGAAGTTCCTGCACAGCGACTCAAGGGCAGCTGCA 301
Db 2451 GCAACATATATTTAGCTAGATTGAAATGGTTAGCTAAAGTTATTCACACTGTGTTA 2392
QY 302 GACTTCTCTGATCACTCCAGTCTGAGCTTCTCTCACCAGGATTCCTTACATTTCATCGCC 361
Db 2391 GATAAATTTATTGAGCCATTTATTATCAATTTTATTACAGCATTTATAACATTTTATTT 2332
QY 362 ATTGGCCCGCAGCAATGCGTGGGTGGCGGATGCTCGGCACACGCTCTACAGGACTTTAT 421
Db 2331 GTAGGGCCCTGTCACCTCGTCAATTAGGTTATTGTTTATCTGATGGATTGACTTGGTTATAT 2272
QY 422 GATTTGCGTGGTCCAGTTCGCGGCTGCTCTTCGCTCTGGTCTACTCACCAGTCTGATC 481
Db 2271 GAATTTGGTGGAGCTATTGCTGACTTATATTTGGTTTATATATGCGCAATCGTCAAT 2212
QY 482 ACTGGTCTGCACCAAGTCTTCCCGCCAAATGAGCTG-----GAGCTGTTTAAAC 529
Db 2211 ACAGGAATGCACCATAGCTTTATTGCAATTGAAACGACATTAATGCTGATGCGACTAAA 2152
QY 530 CAGGTTGGATTCCTTCATCTTCGCAACGGCATCTATGGCTAATATCGCCAGGTGGCGCA 589
Db 2151 ACAGTGGTTCATTTATTTCCCAATCGCAACGATGTCAAATATTCACAAGGTGGTGA 2092
QY 590 TGTTCGCGAGTGTTC---TCCTGGCGAAGAGTGAAGCTCAAGGGCCTTCAGGTGCT 646
Db 2091 GCTTTAGCTGCATCTCTTTATCATTAAGCAAAATAAAAGGTGTGTGCTTCGCG 2032
QY 647 TCAGTGTCTCCGCTCTTCTTGGTATTACGGACCTTCGATCTTCGCTGTGAACCTTCGC 706
Db 2031 CGGGTATTTCAGCTTTACTAGGAATTCAGAACACAGCATGTTTGGTGTCAATCTTAAA 1972
QY 707 CTGCGTGGCGGCTTCCTTCATCGGTATCGGTACCGCAGCTATCGGTGGCGCTTTGATGCA 766
Db 1971 TTGAGATATCCATTTTATAGGTGCTGTTGCAGGATCAGGTATAGGTGCGGCTTATATTTCA 1912
QY 767 CTCTTTAATATCAAGCAGCTTCGTTGGCGGCTGCGAGTTTCTTGGGTGTGTTCTTAT 826
Db 1911 TTCTTCAAAGTAAAGCGATAGCGCTTGTGTACAGCTGGATTACCTGGATTATATCTATA 1852
QY 827 GATGCTCCAGA 837
Db 1851 AATCCTACACA 1841

RESULT 9
ABK73538
ID ABK73538 standard; DNA; 975 BP.
XX AC ABK73538;
XX DT 13-AUG-2002 (first entry)
XX DE Bacillus licheniformis genomic sequence tag (GST) #829.
XX KW Differential gene expression; genomic sequenced tag; GST;
KW altered culture condition; environmental stress;
KW physiological provocation; ds.
XX OS Bacillus licheniformis.
XX PN WO200229113-A2.
XX PD 11-APR-2002.
XX PF 05-OCT-2001; 2001WO-US031437.

CC used for diagnostics. The protein is also used in an assay for enzymatic studies and as a target for antibiotics. This sequence represents one of the novel *S. aureus* genes of the invention

XX
SQ Sequence 861 BP; 259 A; 136 C; 171 G; 295 T; 0 U; 0 Other;

Query Match 11.4%; Score 174.6; DB 8; Length 861;
Best Local Similarity 54.2%; Pred. No. 5.2e-41;
Matches 407; Conservative 0; Mismatches 329; Indels 15; Gaps 2;

QY 167 ATGCCAATGTCCTGTTGGTTAGATGTTGCCAAGCCGTTACAGGCGACCGTG 226

Db 70 ATTCCATATGGATGTTTGGTTTGGTATTAATCAAGTAGGTTATCAAGACAAAGTG 129

QY 227 CTTCTGTGCTGTGTTCTTCGATCTGGCAACGATCGAGAAGTTCCTGCAACAGCGA 286

Db 130 TTACCTATGCTTAGCAGCTATATCTTAGCTCAATTTGAAAAGGTTTACCAAGTT 189

QY 287 CTCACGGCACTGCAAGCTTCCTGATCACTCCAGTGTGACGTTGCTTCACCGGATTC 346

Db 190 ATTCCAACGGTGTAGATAAATTTGTTAACACCAATCTGTATCTATTTTATAACAGCATTT 249

QY 347 CTTACATTCATGCCAATTCGCCAGCAATGCGCTGGTGGCGATGCTGGGCACACGCT 406

Db 250 CTPAACATTTTCATTTGTAGTCCAACTCACTCGCAATTAGTTACTGGTTATCAGATGGT 309

QY 407 CTCACGGGACTTATGATTTCCGTGTGCCAGTCCGGGCTCTGCTCTTCGGTCTGGTCTAC 466

Db 310 TTAACATGCTTTATGAATTTGGTGTGCAATTTGGTGAATTAATTTCCGATTTGTAT 369

QY 467 TCACCAATGTCATCACTGCTGTCGACAGTCCTT-----CCGCGCAATTTAG 514

Db 370 GCTCCGATTTTATTAACAGGTATGCATCATAGCTTTATAGCTGTAGAAACGACATTAAT 429

QY 515 CTGGAGCTCTTAAACAGAGTGTGATCCTTCATCTTCGCAACGCACTATGGCTAATATC 574

Db 430 GCAGATGCCACTAAACGGGTGGATCATTTAATTCGCCAATGCTCAATGTT 489

QY 575 GCCAGGGTGGCGATGTTTGGCAGTGTCTTCCTGGCGAAG---AGTGAAGAGCTCAAG 631

Db 490 GCACAGGTGGTGCAGCAATTCAGCGTTCTTTATTAACAAATTAAGAGTTAA 549

QY 632 GGCCTTGAGGTGCTTCAGTGTCTCCGCTGTTCTTGGTATTAACGAGCTCGGATCTTC 691

Db 550 GGTGTGGCATCTGCCGCAAGTATTTCCAGCAATTTAGTGTATTAACAGAACCGGCTATGTT 609

QY 692 GGTGTGAACCTTCGCTGGCTGGCGTCTTCATCGGTATCGGTACCGCAGCTATCGGT 751

Db 610 GGTGTAACTTAAACTAAGATATCCATTTATGGCGCTATCGTTGGATCAGGTATGTT 669

QY 752 GCGCTTTGATGCACTTTTAAATCAAGGAGTTGCGTGGCGCTCGAGTTCTTTG 811

Db 670 TCAGCATATATGCTTCTTCAAGGTTAAGCAATCGCATTAGGAATCTGCTGATGCCA 729

QY 812 GGTGTGTTTCTATGATGCTCCAGATGATGTCATGTTCTTGGTGTGTCAGTTGTTACC 871

Db 730 GGAATTAATTTCAATCAATCCAGTATCGAGGATGTTACACTACTTTTGTGTATGACA 789

QY 872 TTCTTCATCGCATTCGGCGCAGGATTTGCTT 902

Db 790 ATATCATTCATCTGCTATTAACAGTTACTT 820

RESULT 11

ID ADH84532

XX ADH84532 standard; DNA; 1491 BP.

XX AC ADH84532;

XX AC ADH84532;

DT 22-APR-2004 (first entry)

DE Enterococcus faecalis polynucleotide #2417.

XX

KW Enterococcus faecalis infection; transcription regulatory element;
KW antibacterial; gene; ds.

XX Enterococcus faecalis.

OS US6617156-B1.

FN 09-SEP-2003.

XX 13-AUG-1998; 98US-00134000.

XX 15-AUG-1997; 97US-0055778P.

XX (DOUC/) DOUCETTE-STAMM L A.

XX (BUSH/) BUSH D.

PA Doucette-Stamm LA, Bush D;

XX WPI; 2003-895394/82.

DR P-PSDB; ADH87937.

XX New nucleic acid comprising a sequence encoding an Enterococcus faecalis polypeptide, useful for preparing a composition for diagnosing or treating *E. faecalis* infection.

PT Disclosure; SEQ ID NO 2417; 193pp; English.

XX The invention relates to Enterococcus faecalis polynucleotides and polypeptides. The invention also relates to a recombinant expression vector comprising a polynucleotide operably linked to a transcription regulatory element, a cell comprising a recombinant vector, a method for producing an *E. faecalis* polypeptide, an isolated nucleic acid comprising a sequence not given in the specification, a recombinant vector comprising the nucleic acid and a cell comprising the recombinant vector. The polynucleotides can be used to detect the presence of *E. faecalis* in a sample. The sequences are useful for preparing a composition for diagnosing or treating Enterococcus faecalis infection. This sequence represents an *E. faecalis* polynucleotide of the invention.

CC Sequence 1491 BP; 457 A; 260 C; 267 G; 507 T; 0 U; 0 Other;

Query Match 11.2%; Score 171.6; DB 10; Length 1491;

Best Local Similarity 47.7%; Pred. No. 5.5e-40;

Matches 596; Conservative 0; Mismatches 624; Indels 30; Gaps 2;

QY 176 TGGTCCCTGTTTGGTTAGATGTTGCCAAGCCGTTACAGGCGACCGTCTCTGTTG 235

Db 181 TGGAAATATTTTGGATACCATGTTGCACAAACAAACTATGCCCTACCAAGTAATTCGGTA 240

QY 236 CTGTTGGTTTCTTGGATCTGGCAACGATCGAGAAGTTCTCGCAAGCGACTCAAGGC 295

Db 241 TTAGCTTCGGTAATCTTTTGTCAATATTTGGAATAATTTTCAATAAAACCTTCTTCA 300

QY 296 ACTGCAGACTTCTCTGATCACTCCAGTGTGCTGCTGCTCACCGGATTCCTTACATTC 355

Db 301 TCTATTGATTTTACATTCACACCAATTAATTCGTAATAAATTACTGGAATCTTAATTC 360

QY 356 ATCGCCATTTGCCAGCAATGCGCTGGGTGGCGGATGCTGGCACACAGGTCTACAGGA 415

Db 361 ACAGTTATTTGGTCCCAATGTTTACTATCTAATGGAATCACAGATGCTATTGTTGG 420

QY 416 CTTTATGATTTTGGTGGTCCAGTGGCGGTCTGCTCTTGGTCTGCTTACTCACAATC 475

Db 421 TTATATATGCAACTGGGTTCAATAGGAATGGGATCTTTTGGTGGAAACATATTCACTCAT 480

QY 476 GTCACTACTGTTGTCACAGTCTTCCGCCAATTTAGCTGAGCTGGAGCTGTT----- 525

Db 481 GTTATGACTGTTCTTCATCAGTCAITTCCTGCAATCGAAACACAGTTACTTTCAGCATGG 540

QY 526 -----TAACCGAGGTGGATTCCTTCACTTCGCAACGGCAATCTATGGCTAATATCGCC 577

Db 541 ACAATGATGATCGGACATGGGATTTTATCTTTGTTAGTTGCTTCCATGGCAATAGTCT 600


```
QY 809 TTGGGTGTTGTTCTTATTTGATGCTCCAGATATGGTCAATGTTCTTGGTGTGTCAGTTGTT 868
Db 2330 TTGAGCTTGCTGGCTTTATTTCTGCGAGACAAAAGATATGACTCACCAGATATTAATCTGGT 2389
QY 869 ACCTTCTTCATCGATTTCGGCGCAGCGATTGCTTATGGCCCTTTACTTGGTTCGCCGCAAC 928
Db 2390 GCAATTTGGTCCGGAATTCCTTATCAATGCGTTTGTCTTAACGTTTGTCTTACGTT-- 2447
QY 929 GGCAGCATTTGATCCAGATGCAACCGTGTCTCCAGTGCCTCAGGAACGACCAAGCCGAA 988
Db 2448 -----TTGAAGTCAACCTATATCCAGAAACAGCAACTGAAAACACAGAACTGATAAG 2500
QY 989 GCGAAGACCCCGCAGATTTTCAAGATTCACCATATCCAGGCA---CCTTTGACC 1045
Db 2501 ATGGTGGCACCTGTAAAAACGAATCAAGAGACAAAATTTATTTAGCAAGTCCACTTCAA 2560
QY 1046 GTGGAAGCTATTGCACTGAGCAGCGTCAGCGATGCCATGTTGCCAGCGGAAAGCTTGGC 1105
Db 2561 GTGGAATTTTACCGCTAGAAAAGTACAAGACCCCTGTTTTCCTTCAGGTGCTTTAGGA 2620
QY 1106 TCGGGGTTGCCATGCTCCCAACCAAGGGGCGAGTTAGTTTCTCCGTTGAGTGAAAGATT 1165
Db 2621 AAAGGTGTTGCAATTGAGCGGACTGAGGCAAACTGTATGCCACCCGAGATGTGAAATC 2680
QY 1166 GTGTGGGATTCCTATCTGGCCATGCTTTTCGCGAGTTTCGCAACCAAGCTGAGGATGTTCC 1225
Db 2681 ACCACATTTATTTCCGACAGGACATGCTGTTGGCTTGACGACACAGAGGCG----- 2731
QY 1226 AATGTGATATCTTGATGCACATTTGTTTCGACAGATGTAACCTCAACGCGACGACTTT 1285
Db 2732 ----GTTGAATTAATGATGATTTGGCATGATGACGTCGAATAGATGTAAGGCTTT 2788
QY 1286 AAACCGCTGAAGACAGGCGGCGATGAAGTCAAGCAGGGGAGCTGTGTGTAATTCGAT 1345
Db 2789 GAATTATCAGTGAAACAAGTGATTTCTGTAAAGAGGAGATTTGCTAGTACTTTTGTAT 2848
QY 1346 ATTGATGCCATTAAGCTGACGTTATGAGGTAAACACGCCGATTTGTTTTCGAATTAC 1405
Db 2849 ATTGCTGCCATTAAGAGAGCTGTTATTCGGTAGTTACACCGATTTGTTGTAACGATACG 2908
QY 1406 AA 1407
Db 2909 AA 2910

RESULT 13
ABS98763
ID ABS98763 standard; DNA; 5840 BP.
XX ABS98763;
AC ABS98763;
DT 18-DEC-2002 (first entry)
XX Enterococcus faecalis contig sequence #31.
XX
KW Computer readable medium; Enterococcus faecalis; microbe; growth;
KW pathogenicity; vaccine; resistance; Enterococcal infection; commercial;
KW therapeutic; industrial; fermenting; sugar source; metabolite; vaccine;
KW biochip technology; antibacterial; modulator of nucleic acid expression;
KW contig; ds.
XX
OS Enterococcus faecalis.
XX
FN US2002120116-A1.
XX
PD 29-AUG-2002.
XX
PF 04-MAY-1998; 98US-00070927.
XX
PR 04-MAY-1998; 98US-00070927.
XX
PA (KUNSCH C A.
PA (DILLON P J.
```

```
PA (BARA/) BARASH S.
XX Kunsch CA, Dillon PJ, Barash S;
XX WPI; 2002-750065/81.
XX
PT Computer readable medium having recorded on it a Enterococcus faecalis
PT nucleotide sequence useful for detecting diseases related to Enterococcus
PT infections in animals.
XX
XX Claim 1; Page; 119pp; English.
XX
XX The present invention relates to a new computer readable medium with an
XX Enterococcus faecalis nucleotide sequence. The invention is useful to
XX diagnose the presence of E.faecalis in a sample or determining the
XX presence of a specific microbe in a sample. The invention is also useful
XX for modulating the growth or pathogenicity of E.faecalis, in a vaccine to
XX confer resistance to Enterococcal infection, for commercial, therapeutic
XX and industrial purposes, and for fermenting a particular sugar source or
XX to produce a particular metabolite. The invention is useful for detecting
XX diseases related to Enterococcus infections in animals, and for detecting
XX E.faecalis using biochip technology. The present nucleic acid sequence
XX represents an Enterococcus faecalis contig DNA sequence of the invention.
XX Note: The sequence data for this patent did not form part of the printed
XX specification but was obtained in electronic format directly from USPTO
XX at http.sequedata.uspto.gov
XX
SQ Sequence 5840 BP; 1848 A; 944 C; 1256 G; 1779 T; 0 U; 13 Other;
Query Match 9.9%; Score 151.2; DB 6; Length 5840;
Best Local Similarity 48.5%; Pred. No. 1.4e-33;
Matches 525; Conservative 0; Mismatches 533; Indels 24; Gaps 3;
QY 329 TTGCTGCTCAGCGATTCTTATCATTCATTCGCGCATTCGGCCAGCAATGCGTGGTGGGC 388
Db 1850 TTCCCTTATTAATGCGACCTGTTACCTTCTTAGCTTATTTGGTCTTATCGGGACGGTCAATCGGC 1909
QY 389 GATGTGCTGGCACACGGTCTACAGGACCTTATGATTTTCGGTGGTCCACTCGCGGCTCTG 448
Db 1910 GATGTGTTGGCAAGAGATACAACGCAATTTACGCTTTTAGTCCAAATATACGAGGTTA 1969
QY 449 CTCCTTGGGTCTGCTTACTCACCATTCGTCATCTGCTGCACCGAGTCCTTCCCGCCA 508
Db 1970 TTAATGGGTTCGTTGGCAAGTCTTGGTAATGTTGGTATGATTTGGGCTTTGTACCA 2029
QY 509 ATTGAGCTGGAGCTGTTTAAACCAAGGTGGATCCTTCATCTTCGCAACGCACTATATGGCT 568
Db 2030 ATTATGATGTTAAACTTAAACAAGGTGGCGATACGATGGTACCGATGTTATTACCGCC 2089
QY 569 AATATCGCCCGGGTGGCGCATGTTTGGCAGTGTCTTCTCGCGAGAGTGAAGAGTCC 628
Db 2090 GTTATGTCACAAAGCGGGGCTGCTTTAGCTGTCTTTTCTTAAACAAAATGTGAACATA 2149
QY 629 AAGGGCTTGCAAGTGTCTTCAGGTCTCTCCGCTGTTCTTGGTATTAACGAGGCTCGATC 688
Db 2150 AAAGGTTTGGCTTCTCTTCAAGTATACGACTATTTTGGATTTACTGAAACCACTGTA 2209
QY 689 TTCGGTGTGAACCTTCGCGCTGGCTGGCTGTTCTTCATCGGTATCGGTACCGAGCTATC 748
Db 2210 TATGCGTGTGACTTTTACCATTGAAAAACCAATTTATTCAGCTTGTATTGTTGGTGGGTATC 2269
QY 749 GGTGGCGCTTGTGATTTGCACTCTTTAATATCAAGGAGTGTGGTGGGCGCTGCAGTTTC 808
Db 2270 GGTGGTGCATTTGGGCTATGAATCAGTGAACACTTTTACGTTTGGCTTGGTTAGTATG 2329
QY 809 TTGGGTGTTGTTTCTTATTTGATGCTCCAGATATGGTCAATGTTCTTGGTGTGTCAGTTGTT 868
Db 2330 TTGAGCTTGCCTGGCTTTATTTCTCGACAGACAAAGATACTGCAACCGATGATGTTAGT 2389
QY 869 ACCTTCTTCATCGCATTCGCGCAGCGGATTCGTTATGTCCTTTACTTGGTTGCGCGCAAC 928
Db 2390 GCAATTTGGTCCGGAATTTGCCTTTATCATTCGCTTTGTCTTAAACGTTTGTCTTACGTT-- 2447
```

QY 929 GGCAGCATTTGATCCAGATGCAACCGCTGCTCCAGTGCCTGCAGGAACGACCAAGCCGAA 988
Db 2448 -----TTGAGATCACTTAATCCAGAACAGCACTGAAAAACAGAACTGATAAG 2500
QY 989 GCAGAAGCACCACCGCAGAAATTTTCAAAACGATTCACCATATCCAGGCA---CCTTTGACC 1045
Db 2501 ATGCTGGCACCTGTAAAAACGAATCAAGAAGACAAAATTAATTTAGCAAGTCCACTTCAA 2560
QY 1046 GGTGAAGCTATTGACTGAGCAGGCTGAGCATGCTGTTTCCAGCGGAAAGCTTGGC 1105
Db 2561 GGTGAAATTTTACCGCTAGAAAAAGTACAAGACCCCTGTTTGTCTCAGGTGCTTTAGGA 2620
QY 1106 TCGGCGCTGCTGCTGCTCCCAACCAAGGGGAGTGTAGTTTCTCCGCTGAGTGGAAAGATT 1165
Db 2621 AAAGGTGTTGCAATTGAGCCGACTGAAGGCAAACTGTATGACCCGCGAGATGGTCAATC 2680
QY 1166 GTGTGGCAATCCCATCTGGCCATGCTTTCGACGCTGCAACAAAGGCTGAGATGGTTCC 1225
Db 2681 ACCACATTTATCCGACAGGACATGCTGTGGCTTGACGACAAACAGAGGGC----- 2731
QY 1226 AATGTGATATCTTGATGACATTTGGTTTCACACAGTAACCTCAACGGCAGCACCTTT 1285
Db 2732 ---GTTGAAATTAATTAATGCAATTTGGCATGATACGGTTCGAATTAGATGGTAAAGGCTTT 2788
QY 1286 AACCGCTGAAGAAGCAGCGGATGAAGTCAAAGCAGGGAGCTGCTGTGTGAATTCGAT 1345
Db 2789 GAATATCAGTGAACAAGGTAATCTGTAAAGAGAGATTGCTAGTACTTTTGTAT 2848
QY 1346 ATTGATGCCATTAAAGCTGCGAGTTATGAGGTAACCAACCGCATTTGTTTTCGAATTAC 1405
Db 2849 ATTGCTGCCATTAAAGAGCTGTTATCCGGTAGTTTACACCGATTGTGTAAACGAATACG 2908
QY 1406 AA 1407
Db 2909 AA 2910

RESULT 14
ABD00428/c
ID ABD00428 standard; DNA; 1386 BP.
XX AC
XX ABD00428;
XX DT 29-JUL-2004 (first entry)
XX DE Klebsiella pneumoniae polynucleotide seqid 6203.
XX KW Recombinant expression vector; transcription regulatory element;
XX OS Klebsiella pneumoniae protein; antibacterial; Vaccine; gene; ds.
XX PN Klebsiella pneumoniae.
XX US6610836-B1.
XX PD 26-AUG-2003.
XX PF 27-JAN-2000; 2000US-00489039.
XX PR 29-JAN-1999; 99US-0117747P.
XX PA (GENO-) GENOME THERAPEUTICS CORP.
XX PI Breton GL, Osborne M;
XX WPI; 2003-895346/82.
XX DR P-P8DB; ABO66857.
XX PT New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
XX PT preparing a vaccine composition against Klebsiella pneumoniae.
XX PS Disclosure; SEQ ID NO 6203; 932pp; English.
XX XX The invention describes a new isolated nucleic acid encoding a Klebsiella

CC pneumoniae polypeptide. Also described are: a recombinant expression
CC vector comprising the nucleic acid, operably linked to a transcription
CC regulatory element; and a cell comprising the recombinant expression
CC vector. The nucleic acid is useful for preparing a vaccine composition
CC against Klebsiella pneumoniae. This sequence encodes a Klebsiella
CC pneumoniae polypeptide of the invention
XX
SQ Sequence 1386 BP; 300 A; 415 C; 437 G; 234 T; 0 U; 0 Other;
Query Match 9.4%; Score 143.6; DB 11; Length 1386;
Best Local Similarity 51.3%; Pred. No. 1.2e-31;
Matches 366; Conservative 0; Mismatches 339; Indels 9; Gaps 1;
QY 183 TGTGTGTTTGTAGATGTTGCCAACCGGTTTACAGGGACCGTCTCTCTGCTGTGCTGG 242
Db 794 TCTTCGCATCGAAGTGGCGATGATCGCTACCAAGGACCGTCTTCGCGGTGCTGTGG 735
QY 243 TTTCTTGATTTCTGGCAACGATCGAGAAGTCTTCGCAACAGGACTCAAGGGCACTGCAG 302
Db 734 CGGTGTGTTTATGAGCATGGTTCGAGAAGCGGCTCGCGCGGTGATCCCTGACGGCTGG 675
QY 303 ACTTCTGATCACCTCCAGTGTGAGCTTGTGCTCACCGGATTCCTTACATTCATCGCA 362
Db 674 ACCTGATCCTCCTCGTTCCTGAGCGTGAATATCTCGGCTTTATCGCCCTGCTGCTGA 615
QY 363 TTGCCCCAGCAATGCGCTGGGTGGCGGATGTCTGGCACACGGTCTACAGGACATTTATG 422
Db 614 TCGCCCCGCGCTCGCGCGCTCGGCGACGCAATTTCTGTTATCTCAGCAGCTTATCA 555
QY 423 ATTTGGTGTCTCAGTTCGCGGCTGCTCTTCTGGTCTGCTACTCACCATTCGTCATCA 482
Db 554 GCCAGCGCGGTGGGTGGCGGCTGCTGTTTGGCGGCTCTATTTCGGTGTGCTGTTATTA 495
QY 483 CTGCTGTCACCACTCTTCCCGCAATTGAGCTCGAGCTGTTTAAACA-----GG 533
Db 494 CGGTATCCATCAGCTTCCATGCAATCGAGGCGGACTGCTGGCAACCCATCGATTG 435
QY 534 GTGGATCCTTCTTCCCAACCGCATCTATGGTAAATATGCCCGAGGTGCGGATGTT 593
Db 434 CGGTCAACTTCTCTGCTGCGGATCTGGCGATGGCCAAAGTCCGCCAGGCGCGCTGCT 375
QY 594 TGGCAGTGTCTTCTGCGGAGAGTGAAGAGCTCAAGGGCTTCAGGTGCTTCAGGNG 653
Db 374 TTTGGTGTGTTTAAACCAAGATGCCAAATAAAGCTATCACCTGCGCGT 315
QY 654 TCTCGCTGTTCTTCTGCTATTACGAGCTGCGCATCTTGGTGTGAACCTTCCCTCGCT 713
Db 314 TTTCCGCGATGCTGGGATCACCGAGGCGGCAATCTTGGGATTAACCTGCGCTTGTGA 255
QY 714 GCGCGTCTTCTATCGGTATCGGTACCGCAGCTATCGGTGGCGCTTTGATTGCACTTTA 773
Db 254 AACCGTTCATCGCGCGCTGTTGGGCGGTGCGCGCGCGCGCTGGGTGTGCTGATGC 195
QY 774 ATATCAAGCAGTTGCTTGGGCGCTGAGGTTCTTCTGGTGTGTTTCTATTATGATGCTC 833
Db 194 ACGTCTATCATGACCGCGGTGGGCTTACCGCGATCCCGGGAATGGGTATCGTGACGCCA 135
QY 834 CAGATATGTCATGTTCTTGGTGTGTCAGTGTGTTTACCTTCTTCATCGCATTCG 887
Db 134 GCTCGTGTGAACTACATATTCGGAATGCGATCCCTTCGCGGTGGCTTCG 81
RESULT 15
ABD00072
ID ABD00072 standard; DNA; 1452 BP.
XX AC
XX ABD00072;
XX DT 29-JUL-2004 (first entry)
XX DE Klebsiella pneumoniae polynucleotide seqid 5847.
XX XX Recombinant expression vector; transcription regulatory element;

Klebsiella pneumoniae protein; antibacterial; Vaccine; gene; ds.

Klebsiella pneumoniae.

US6610836-B1.

26-AUG-2003.

27-JAN-2000; 2000US-00489039.

29-JAN-1999; 99US-0117747P.

(GENO-) GENOME THERAPEUTICS CORP.

Breton GL, Osborne M;

WPI: 2003-895346/82.

P-PSDB; AB066501.

New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for preparing a vaccine composition against Klebsiella pneumoniae.

Disclosure; SEQ ID NO 5847; 932pp; English.

The invention describes a new isolated nucleic acid encoding a Klebsiella pneumoniae polypeptide. Also described are: a recombinant expression vector comprising the nucleic acid, operably linked to a transcription regulatory element; and a cell comprising the recombinant expression vector. The nucleic acid is useful for preparing a vaccine composition against Klebsiella pneumoniae. This sequence encodes a Klebsiella pneumoniae polypeptide of the invention

Sequence 1452 BP; 256 A; 452 C; 429 G; 315 T; 0 U; 0 Other;

Query Match 9.4%; Score 143.6; DB 11; Length 1452;

Best Local Similarity 51.3%; Pred. NO. 1.2e-31;

Matches 366; Conservative 0; Mismatches 339; Indels 9; Gaps 1;

QY	183	TGTTGGTTTAGATGTGCCAAGCGGTACAGGGCACCGTGTCTCTGTGGTGGTG 242
Db	698	TCTTCGGCATCGAAGTGGCGATGATCGGCTACAGGGCACCGTCTTCGCGTGTCTGG 757
QY	243	TTTCTTGGATTCTGGCAAGATCGAGAAGTTCTCGCAAGCGACTCAAGGGCACTGCAG 302
Db	758	CGGTGTGGTTTATGAGCATGTTCGAGAAGCGGTGGCGCGTGAATCCCTGACGGCTGG 817
QY	303	ACTTCTGATCACTCCAGTGTGACGTGTCTGCTCACCGGATTCCTTACATTCGCCA 362
Db	818	ACCTGATCTCACTCCGTTCCCTGACGGTGATTAATCTCCGGCTTTATCGCCCTGCTGA 877
QY	363	TTGGCCCAAGCAATGGCTGGGTGGCGATGTGTGGCACACGGTCTACAGGGACTTTATG 422
Db	878	TCGGCCCGCGCGTGGCGCTCGCGCAGCGGATTTGTTTATCTCAGCAGCTTATCA 937
QY	423	ATTTCGGTGGTCCAGTCGGCGGTCTGCTCTTCGGTCTGTCTACTACCAATCGTCATCA 482
Db	938	GCCAGCGCGGTGTGGTGGGGGCGCTGCTCTTCGGGGCGCTTATTTCGGTGTATGTTA 997
QY	483	CTGGTCTGACCAAGTCTTCGCCCAATAGCTGAGCTGTTAACCA-----GG 533
Db	998	CCGGTATCCATCACAGTTCATGCCATCGATCAGGCGCGGACTGTGGGCACCCCATCGATTG 1057
QY	534	GTGGATCTCTTCAATTCGCAAGCGCATCTATGGCTAATATCGCCAGGGTGGCGATGTT 593
Db	1058	GGGTCAACTTCTCTGTCGGATCTGGGCGATGGCCAACTCGCCAGGGCGGCGCTGCT 1117
QY	594	TGGCAGTGTCTTCTGGCGAAGATGAAGAACTCAAGGGCGCTTCGAGGTGCTCAGGTG 653
Db	1118	TTGCGGTGTGGTTTAAACCAAGATGCCAAATAAAAGCTATACCTCGCGCTGGCGT 1177
QY	654	TCTCCGCTGTTCTTGGTATACGGAGCTCGGATCTTGGTGTGAACCTTCGCCCTGGCT 713
Db	1178	TTTCGGCGATGTGGGGATCACCGAGGGGCAATCTTCGGGATTAACTTCGCTTTGTGA 1237

QY	714	GGCGTTTCTTCATCGGTATCGGTACCGCAGCTATCGGTGGCGCTTTGATTCACCTCTTAA 773
Db	1238	AACCGTTTCATCGCGCGCTGGTGGCGGTGCCCGCGCGCTGGGTGGTGTTCGATGC 1297
QY	774	ATATCAAGGCAGTTGGCTTGGCGCTGCAGGTTTCTTGGGTGTGTTTCTATTGATGCTC 833
Db	1298	ACGTCTACATGACCGCGGTGGGCTGACCGCGATCCCGGGAATGGCTATCGTGCAGGCCA 1357
QY	834	CAGATATGTCATGTTCTTGGTGTGTCAGTGTGTACCTTCTTCATCGCATTCG 887
Db	1358	GTCGCTGCTGAACATACATTACGGAATGGCGATCGCCTTCGCCGTGGCTTCG 1411

Search completed: October 30, 2004, 14:04:57

Job time : 734 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 30, 2004, 13:46:16 ; Search time 154 Seconds
(without alignments)
7047.887 Million cell updates/sec

Title: US-09-604-231-1

Perfect score: 1527

Sequence: 1 cccatggcatcgcgcgtt.....gttgaaccttgatgttgcg 1527

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/prodata/1/ina/5A COMB.seq.*
2: /cgn2_6/prodata/1/ina/5B COMB.seq.*
3: /cgn2_6/prodata/1/ina/6A COMB.seq.*
4: /cgn2_6/prodata/1/ina/6B COMB.seq.*
5: /cgn2_6/prodata/1/ina/6C COMB.seq.*
6: /cgn2_6/prodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	190.6	12.5	2913	4	US-09-710-279-3464, Ap
C 2	171.6	11.2	1491	4	US-09-134-000C-2417, Ap
C 3	143.6	9.4	1386	4	US-09-489-039A-6203, Ap
C 4	143.6	9.4	1452	4	US-09-489-039A-5847, Ap
C 5	137.4	9.0	732	3	US-09-134-001C-1277, Ap
C 6	126.6	8.3	1968	4	US-09-583-110-297, Ap
C 7	126.6	8.3	3895	4	US-08-961-527-201, Ap
C 8	117.6	7.7	1884	4	US-09-583-110-463, Ap
C 9	114.6	7.5	951	4	US-09-134-000C-2077, Ap
C 10	109.6	7.2	3615	1	US-08-920-812-17, Ap
C 11	109.6	7.2	3615	1	US-08-920-827-17, Ap
C 12	109.6	7.2	3615	1	US-08-921-177-17, Ap
C 13	109.6	7.2	3615	1	US-08-362-577C-17, Ap
C 14	109.6	7.2	3615	2	US-08-920-828-17, Ap
C 15	109.4	7.2	1905	4	US-09-543-681A-2847, Ap
C 16	107.8	7.1	9769	4	US-08-961-527-30, Ap
C 17	97.8	6.4	465	2	US-08-673-190A-3, Ap
C 18	97.2	6.4	1971	4	US-09-489-039A-1041, Ap
C 19	90	5.9	1428	4	US-09-489-039A-1806, Ap
C 20	86.2	5.6	30246	4	US-08-956-171E-56, Ap
C 21	86.2	5.6	30246	4	US-08-781-986A-56, Ap
C 22	83.2	5.4	1284	4	US-09-107-532A-1876, Ap
C 23	81	5.3	1887	4	US-09-107-532A-1634, Ap
C 24	76.4	5.0	465	4	US-08-956-171E-1317, Ap
C 25	76.4	5.0	465	4	US-08-781-986A-1317, Ap
C 26	71.6	4.7	357	2	US-08-673-190A-6, Ap
C 27	71	4.6	270	4	US-09-134-000C-2419, Ap

28	70.4	4.6	2013	4	US-09-134-000C-1201	Sequence 1201, Ap
29	67	4.4	2181	4	US-09-583-110-121	Sequence 121, App
30	67	4.4	8494	4	US-08-961-527-163	Sequence 163, App
31	64.4	4.2	1446	4	US-09-543-681A-984	Sequence 984, App
32	60.4	4.0	1437	3	US-09-134-001C-2228	Sequence 2228, Ap
33	60.4	4.0	2550	4	US-09-710-279-4343	Sequence 4343, Ap
34	58.8	3.9	1752	4	US-09-107-532A-3157	Sequence 3157, Ap
35	57.6	3.8	1971	4	US-09-107-532A-1429	Sequence 1429, Ap
C 36	56.8	3.7	29555	4	US-08-956-171E-206	Sequence 206, App
C 37	56.8	3.7	29555	4	US-08-781-986A-206	Sequence 206, App
38	55.8	3.7	1846	4	US-09-634-238-183	Sequence 183, App
39	55.4	3.6	1896	4	US-09-107-532A-248	Sequence 248, App
40	54.6	3.6	591	4	US-09-543-681A-3487	Sequence 3487, Ap
C 41	53.2	3.5	2996	4	US-09-710-279-3809	Sequence 3809, Ap
C 42	53.2	3.5	3031	4	US-09-710-279-4310	Sequence 4310, Ap
C 43	53.2	3.5	3932	4	US-09-710-279-3420	Sequence 3420, Ap
C 44	52.2	3.4	315	4	US-09-134-000C-2078	Sequence 2078, Ap
C 45	51.4	3.4	1839	4	US-09-583-110-250	Sequence 250, App

ALIGNMENTS

RESULT 1

US-09-710-279-3464/c
; Sequence 3464, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PUB480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3464
; LENGTH: 2913
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-09-710-279-3464

Query Match	12.5%	Score	190.6;	DB	4;	Length	2913;
Best Local Similarity	55.0%;	Pred. No.	3.7e-48;				
Matches	468;	Conservative	0;	Mismatches	364;	Indels	19;
Gaps	4;						
QY	5	TGGCATCTCGCGGTCGGTTCCTTGGCGCGCGTATTTGGTATCGCATGCTGTTCCGAGCTTG	124				
Db	2690	TTGCAATGACCTTTTACATATTACCAATACCTTATTGGTTTTAGTCAGCTAAGCGAT	2631				
QY	65	TGGCGGCGCAATGAGTTCCTTGGCGCGCGTATTTGGTATCGCATGCTGTTCCGAGCTTG	124				
Db	2630	TTGGTGGTAATCCTTATTAGTGAGC-TCTAGGTATGATCTTGTTCATCTGATTTG	2572				
QY	125	GTGAACGGGTACGACGTGGCGCCACCATGGCTGGCGGGCGAA---ATGCCAATGTGGTCC	181				
Db	2571	ATGAGTGCATATGATTTTCCAAAAGCTTTAGAAGAAGGAAAAGCTATTCCACACTGGAT	2512				
QY	182	CTGTTTGGTTTATGATTTGCCCAACCGGTTCACAGGGCACCGTGTCTCTGTGTGTTG	241				
Db	2511	GTCTTTGGGTACATATTAAATGAAGTAGTTATCAGGGACAAAGTATTACTATGCTGTA	2452				
QY	242	GTTCCTTGGATTCTCGCAACGATCGAGAAGTTCCTGCAACCGACTCAAGGGCACTGA	301				
Db	2451	GCAATATATTTTATGCTACGATTGAAAATAGTTACGTAAAGTTATTCCAACTGTGTTA	2392				
QY	302	GACTTCTCGTACATCCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	361				
Db	2391	GATAATTTTTCAGCGCATTTATCAATTTTATTACAGCATTTATACATTTTATTTT	2332				

RESULT 3

US-09-489-039A-6203/c
; Sequence 6203, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 6203
; LENGTH: 1386
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-6203

Query Match 9.4%; Score 143.6; DB 4; Length 1386;
Best Local Similarity 51.3%; Pred. No. 7.7e-34;
Matches 366; Conservative 0; Mismatches 339; Indels 9; Gaps 1;

QY	183	TCGTTGGTTTAGATGTTGCCAAGCGGTTACCAAGGCGACCGTGCTTCTGCTGCTGGTGG 242
DB	794	TCCTCGGCATCGAAGTGGCGATGATCGGCTACAGGGCACCGTCTCCCGGTGCTGCTGG 735
QY	243	TTTCTTGGAATTCGGAACGATCGAAGTTCTCTGCAAGCGACTCAAGGGCACTGCGAG 302
DB	734	CGGTGTGGTTTATGAGCATGCTCGAAGCGGCTGCGCGCGTATCCCTGACGCGCTGG 675
QY	303	ACTTCTGATCACTCAGTGTCTGCTGCTGCTACCGGATTCCTTACATTCATCGCCA 362
DB	674	ACCTGATCTCACTCCGTTCCCTGAGCGGTGATTTCTCGCGCTTTATCGCCCTGCTGTA 615
QY	363	TTGGCCCAAGCAATGCGTGGGTGGCGATGTGCTGGCAACAGGTCTACAGGACATTTAG 422
DB	614	TCGGCCCGCGGTGCGCGCTCGGCGAGCGCATTTGTTTATCTCAGCAGCTTATCA 555
QY	423	ATTTGCGTGGTCAAGTGGCGGTGCTCTTCCGTTCTGCTTACTACCAATCGTCATCA 482
DB	554	GCCACGCGCGGTGCGTGGCGGCGCTGCTGTTGCGGCGCTTATTCGGTATCGTTATTA 495
QY	483	CTGCTCTGCAACAGTCTTCCCGCAATGAGCTGAGCTGTTTAAACA-----GG 533
DB	494	CCGGTATCCATCACAGCTTCCATGCAATGAGGCGCGGACTGCTGGGCAACCCATCGATTG 435
QY	534	GTGGATCTTTCATCTTCGCAACGGCATCTATGGCTAATATCCCGAGGTGGCGCATGTT 593
DB	434	CGGTCAACTTCTGCTGCGGATCTGGGCGATGCGCAACGTGCGCCAGGGCGCGCTGCT 375
QY	594	TGGCAGTGTCTTCTGCGGAAGATGAAAGCTCAAGGCGCTTGCAGGTGCTTCAGGTG 653
DB	374	TTGCGGTGTGGTTTAAACCAAGATGCCAAATAAAGCTATACCTCGCGCTCGCGGT 315
QY	654	TCTCGCTGTTCTTGGTATTACGGAGCGCTCGGATCTTCGGTGTGAACCTTCGCTCGCGT 713
DB	314	TTTCGGCGATGCTGGGGATCACCGAGGCGGCAATCTTCGGGATTAACCTGCGCTTTGTA 255
QY	714	GGCGCTTCTTCATCGGTATCGGTACCGAGCTATCGGTGGCGCTTTGATGACATCTTTA 773
DB	254	AACCGTTTCATCGCGGTGTTGGCGGTGCGCGCGGCGCGCTTGGTGGTTCGATGC 195
QY	774	ATATCAAGCGAGTTCGTTGGCGCTGACAGGTTTCTTGGTGTGTTTCTATTTGATGCTC 833
DB	194	ACGTCTACATGACCGCGGTGGCGCTGACCGCGATCCCGGAATGCTATCGTGCAAGGCA 135
QY	834	CAGATATGCTATGTTCTGTTGTTGTCAGTTGTTTACCTTCTTCATCGCATTCG 887
DB	134	GCTCGCTGCTGAACCTACATTAATCGGAATGGGATCGCCTTCGCGGTGSCCTTCG 81

RESULT 4

US-09-489-039A-5847
; Sequence 5847, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 5847
; LENGTH: 1452
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-5847

Query Match 9.4%; Score 143.6; DB 4; Length 1452;
Best Local Similarity 51.3%; Pred. No. 7.9e-34;
Matches 366; Conservative 0; Mismatches 339; Indels 9; Gaps 1;

QY	183	TCGTTGGTTTAGATGTTGCCAAGCGGTTACCAAGGCGACCGTGCTTCTGCTGCTGGTGG 242
DB	698	TCCTCGGCATCGAAGTGGCGATGATCGGCTACCAAGGCGACCGTCTTCGCGGTGCTGCTGG 757
QY	243	TTTCTTGGAATTCGGAACGATCGAAGTTCTCTGCAAGCGACTCAAGGGCACTGCGAG 302
DB	758	CGGTGTGGTTTATGAGCATGCTCGAAGCGGCTGCGCGCGTATCCCTGACGCGCTGG 817
QY	303	ACTTCTGATCACTCAGTGTCTGCTGCTCACCAGATTCCTTACATTCATCGCCA 362
DB	818	ACCTGATCTCACTCCGTTCCCTGACGCGTATCTTCGCGCTTTATCGCCCTGCTGTA 877
QY	363	TTGGCCCAAGCAATGCGTGGGTGGCGATGTGCTGGCAACAGGTCTACAGGACATTTATG 422
DB	878	TCGGCCCGCGGTGCGCGCTCGCGCGATTCGTTTATCTCAGCAGCTTATCA 937
QY	423	ATTTGCGTGGTCCAGTTCGCGGCTGCTCTTCGGTCTGCTGCTACTACCAATCGTCATCA 482
DB	938	GCCACGCGCGGTGCGTGGCGGCGCTGCTGTTGCGGCGCTTATTCGGTATCGTTATTA 997
QY	483	CTGCTCTGCAACAGTCTTCCCGCAATGAGCTGAGCTGTTTAAACA-----GG 533
DB	998	CCGGTATCCATCACAGCTTCCATGCAATGAGGCGCGGACTGCTGGGCAACCCATCGATTG 1057
QY	534	GTGGATCTTTCATCTTCGCAACGGCATCTATGGCTAATATCCCGAGGTGGCGCATGTT 593
DB	1058	CGGTCAACTTCTGCTGCGGATCTGGGCGATGCGCAACGTGCGCCAGGGCGCGCTGCT 1117
QY	594	TGGCAGTGTCTTCTGCGGAAGATGAAAGCTCAAGGCGCTTGCAGGTGCTTCAGGTG 653
DB	1118	TTGCGGTGTGGTTTAAACCAAGATGCCAAATAAAGCTATACCTCGCGCTCGCGGT 1177
QY	654	TCTCGCTGTTCTTGGTATTACGGAGCGCTCGGATCTTCGGTGTGAACCTTCGCTCGCGT 713
DB	1178	TTTCGGCGATGCTGGGGATCACCGAGGCGGCAATCTTCGGGATTAACCTGCGCTTTGTA 1237
QY	714	GGCGCTTCTTCATCGGTATCGGTACCGAGCTATCGGTGGCGCTTTGATGACATCTTTTA 773
DB	1238	AACCGTTTCATCGCGCGCTGTTGGCGGTGCGCGCGGCGCGCTTGGTGGTTCGATGC 1297
QY	774	ATATCAAGCGAGTTCGTTGGCGCTGCGAGTTCCTTCGGGTGTTGTTTCTATTTGATGCTC 833
DB	1298	ACGTCTACATGACCGCGGTGGCGCTGACCGCGATCCCGGAATGCTATCGTGCAAGGCA 1357
QY	834	CAGATATGCTATGTTCTTGGTGTGTCAGTTGTTTACCTTCTTCATCGCATTCG 887
DB	1358	GCTCGCTGCTGAACCTACATTAATCGGAATGGGATCGCCTTCGCGGTGSCCTTCG 1411

RESULT 5

US-09-134-001C-1277
; Sequence 1277, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134.001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 1277
; LENGTH: 732
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (20), (22), (32)
; OTHER INFORMATION: Identity of nucleotide at the above locations are unknown.
US-09-134-001C-1277

Query Match 9.0%; Score 137.4; DB 3; Length 732;
Best Local Similarity 56.4%; Pred. No. 4.2e-32;
Matches 305; Conservative 0; Mismatches 221; Indels 15; Gaps 2;

QY 312 TCATCCAGTGTGAGTGTGCTCACCGGATTCCTTACATTCATCGCATGCGCCAG 371
DB 80 TGACGCCATTAATATCAATTTTATTACAGCATTTATAACATTTTATTGTAGGCGCTG 139

QY 372 CAATCCGCTGGTGGCGATGTGCGCACAGCGTCTACAGGACTTTATGATTTTCGGTG 431
DB 140 TCATCGTCAATTAGTTATTTGGTTATCTGATGGATTGACTTGGTTATATGAATTTGGTG 199

QY 432 GTCCAGTCCGGGCTGTCTTTCTGCTGCTGCTTCTACTCCAAATGCTATCTGCTGCTG 491
DB 200 GAGCTATTGGTGAATTAATTTGTTTATTATATATGCGCAATCGTCAATACAGGAATGC 259

QY 492 ACCAGTCTTCCGCGCAATTTGAGCTG-----GAGCTGTTTAAACAGGTTGGAT 539
DB 260 ACCATATCTTTATTCAGTTGAACGACATTAATTTGCTGATCGCACTAAACAGGTGGTT 319

QY 540 CTTTCATCTTCGCAACGGCATCTATGGCTTAATATCGCCAGGCTGGCGCATGTTTGGCAG 599
DB 320 GATTATCTTCCATTCGCAACGATGCAAAATATTCGCAAGGTGGTGGAGCTTTAGCTG 379

QY 600 TGTTCT---TCTGGCGAAGAGTGAAGCTCAAGGCGCTTGCAGGTGCTTCAAGTGCT 656
DB 380 CATTCTTATCATTAAGCAAAATAAAAAATTAAAAAGTCTTGCTTCCGCGCGGGTATTT 439

QY 657 CGCTGTTCTGTTATTAAGGACCTGCGATCTTGGTGTGAACCTTCCGCTGCGCTGAC 716
DB 440 CAGCTTTACTAGGAATACAGAACCGCAATGTTTGGTGTCAATCTTAAATTTAGATATC 499

QY 717 CGTTCTTCACTCGGTATCGGTACCGAGCTATCGGTGGCGCTTTGATTTGACACTTTTAAATA 776
DB 500 CATTATAGGTGCTGTGTCAGGATCAGGTATAGGTGCGGCTTATATTTCACTTCAAAG 559

QY 777 TCAGGCAAGTGGGTGGCGCTGCAAGTTTCTTGGGTGTTGTTCTTATTTGATGCTCCAG 836
DB 560 TAAAGCGATAGCGCTTGTACAGCTGGATTACCTGGATTTATCTATAAATTCCTACAC 619

QY 837 A 837
DB 620 A 620

RESULT 6

US-09-583-110-297
; Sequence 297, Application US/09583110
; Patent No. 669703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 297
; LENGTH: 1968
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-297

Query Match 8.3%; Score 126.6; DB 4; Length 1968;
Best Local Similarity 47.0%; Pred. No. 1.7e-28;
Matches 581; Conservative 0; Mismatches 629; Indels 27; Gaps 5;

QY 174 TGTGGTCCCTGTTTGGTTTAGATGTTGCCAAGCGGTTACAGGGCACCGTGTCTCTG 233
DB 662 TTGGAATTTTGGTTATTTTACTGTTTAATGTTATGTTTACCAAGCCCAAGTTATCCAG 721

QY 234 TGCTGGTGGTTTCTTGGATTCTGGCAACGATCGAGAAGTTCTGTGCACAGCGACTCAAG 293
DB 722 CTTGCTTGCAGTTTGAATCTGTCTTATCTGAAATTTTGGCGCAACATATCCAG 781

QY 294 GCATGCGAGACTTCTGTATCATCTCAGTGTGACGTTGCTGTGTCTACCGGATTCCTTACAT 353
DB 782 AAGTCATTTCATGATATTTTGTACCTTTTGTCAATTCATTCAGCCTTGATTTTGGCTC 841

QY 354 TCATCGCATTTGGCCCGCCAGCAATGCGCTGGTGGCGGATGTGTCGACACAGGCTCTACAG 413
DB 842 ATACTGTCTTGGGCGCAATCGGTTGGACAAATGGAACAAGGACTTTCATCAGTTGTTAG 901

QY 414 GACTTT---ATGATTTTCGGTGGTCCAGTCGGCGGTCTGCTCTTCGGTCTGCTACTAC 470
DB 902 CAGGTTTAAACGGTCTCTTAAATGCTCTTCGGTGCAATTTTGGTGGCTCTACGCTC 961

QY 471 CAATCGTCATCATCTGGTCTGCAACAGTCTCTCCCGCCAAATTCAGCTGGAGCTGTTT---A 527
DB 962 CATTGTTCATCACAGGCTCTGCACCATGACCAATGCCATTCATACACAATTTGATGCGG 1021

QY 528 ACCAGGTGGATCCTTTCATCTTCGCAACGCACTATGCTATATATGCGCCAGGGTGGCG 587
DB 1022 ATGCTGGTGGCACTGCGCTCTGGCTATGATTTGCTCTTCTTAAATTTGCTCAGGCTCAG 1081

QY 588 CATGTTTGGCAGTGTCTTCTTCGGCGAAGAGTGAAGAGCTCAAGGSC---TTGCAAGGTG 644
DB 1082 CCGTGTTCCTTATTTATTCATGTCATCGCATCGCCATGATGAGTGTGAGGCTCAGGTTCACTTC 1141

QY 645 CTTTCAGGTGCTCCGCTGTTCTTGGTATTTACGAGGCTCGGATCTTCGGTGTGAACCTTC 704
DB 1142 CTCGAACCAATTTTCAGCCTATCTCGGTGTACGAACCAAGCTCTTTTGGGGTTAAGTAA 1201

QY 705 GCCTCGCTGGCGGTTCTTTCATCGGTATCGGTACCGAGCTATCGGTGGCGCTTTGATTTG 764
DB 1202 AATATATTTATCAATTTGTTGCTGGATGACTGTTTCAGCCCTTCAGGCAATGTTATCCG 1261

QY 765 CACTCTTTAATATCAAGGCAAGTTGGTGGCGCTGCAAGTTTCTTGGGTGTTGTTCTA 824
DB 1262 TTACTTTTAACTGTAAGTGGGCTTCTTATTTGTTATCGGTGGTTTGGCAGGATTTCTCTCA 1321

QY 825 TTGATGCTCCAGATATGTCATGTTCTTGTGTGTGTCAGTTGTTTACCTTCTTTCATGCGAT 884

Db	1322	TTCAACTCAATACATGCTGCCATTTGCGAACAATGCTAGTTGCGAATGTTGTTCCAA	1381
Qy	885	TCGCGCGACGGAATGCTTATGGCTTTACTTTGTTTCGCCGCAACGGCAGCAATGATCCAG	944
Db	1382	TGCTC-----TTGACTTCTCTTTTCGCAAGGCTGGTCTCTTCACAAAATAGAGGGCG	1435
Qy	945	ATGCAACGGCTGCTCCAGTGCCTTCAGGACGACCAAAAGCGGAAGCAGAAGCACC CGCAG	1004
Db	1436	ATACGAACCTTGCAGCGAGAATTGGTCTCAAGAAGAAGACAGAAATTTGTGAGCCATGAAC	1495
Qy	1005	AAATTTTCAAAACGATTCACCATCATCAGGACCTTTGACCGGTGAAGCTATTGCACTGA	1064
Db	1496	CAGTAGAACCTTCTCGGTAGAAATATCAGCCCACTAACTGGCCCAAGTGAAGAATTTGA	1555
Qy	1065	GCAGCGTCAGCGATGCCATGTTTGCCAGCGGAAAGCTTGGCTCGGGCGTTGCCATCGTCC	1124
Db	1556	GTCAAGCGACGGATCCCTGTTTTTGATCAGGTGCATGGGCGCAAGGTCCTGGTCATTGAAC	1615
Qy	1125	CAACCAAGGGGCACTGATTTCTCCGGTGAGTGGAAGATGTGTGGCAATGCCATCTG	1184
Db	1616	CAAGCAGGTGAGTTGACCTCTCCAGTTAATGGAAACAGTGACGGTTCCTTTCCCTACCA	1675
Qy	1185	GCCATGCTTTCCGAGTTTCGACCAAGGCTGAGATGTTCCAAATGTGGATATCTTGATGC	1244
Db	1676	AGCATGCCATCGGCATTGTCTCTGACGAGGGAGTTGAAT-----TGCTCATCC	1723
Qy	1245	ACATTGGTTTCGACACAGTAAACCTCAACGGCAGCAGCTTTAAACCGCTGAAGAAGCAGG	1304
Db	1724	ACATCGGTATGGATACAGTAGGCTCTTGATCGCAAAAGGTTTTGAAAGTCTTTGTAGTCCAAG	1783
Qy	1305	GCGATGAAGTCAAGCAGGGGAGCTGTGTGTAATTCGATATTGTATGCCATTAAAGGCTG	1364
Db	1784	GAGATCAGTTATAGTTGGCCAGCACTGATTCGTTTTGATATGATGATGCTTTAAGGCTG	1843
Qy	1365	CAGGTTATAGGTAAACACCGCGATTGTTGTTTCGAA	1401
Db	1844	CAGGTCCTGGTGACAGAAACCTCCTGTTATCATCACCA	1880

RESULT 7

```

US-08-961-527-201/c
; Sequence 201, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504

```


GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
FILE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 032796-032
CURRENT APPLICATION NUMBER: US/09/134,000C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/055,778
PRIOR FILING DATE: 1997-08-15
NUMBER OF SEQ ID NOS: 6812
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2077
LENGTH: 951
TYPE: DNA
ORGANISM: Enterococcus faecalis
US-09-134-000C-2077

Query Match 7.5%; Score 114.6; DB 4; Length 951;
Best Local Similarity 48.4%; Pred. No. 5.5e-25;
Matches 429; Conservative 0; Mismatches 434; Indels 24; Gaps 3;

QY 524 TTTAAACAGGCTGGATCCCTTCATCTTCGCAACCGCATCTATGGCTAATATCGCCAGGCT 583
DB 10 TTAACACAGAGTGGCATAGGATCGTACCGATGTTATACAGCCGTTATTCACAGGC 69
QY 584 GCGGCAAGTTTGGCAGTGTCTTCTCTGCGCAAGAGTGAAAAGCTCAAGGCGCTTCGAGGT 643
DB 70 GGGGCTGCTTTAGCTCTCTTTTCTTCCCTAACAAAAATGTGAACTAAAGSTTGGCTTG 129
QY 644 GCTTCAGGTGTCCTCGCTCTCTTGTGTATACGAGCCCTGCGATCTTCGGTGTGAACCTT 703
DB 130 TCTTCAAGTATTACGACTATTTTGGAAATTAAGTAACTGATATATGTCGCTGACTTTA 189
QY 704 CGCCTGCGTGGCGGCTTCCTATCGGTATCGGTACCGCAGCTATCGTGGCGCTTTGATT 763
DB 190 CCATTGAAAAACCAATTAATGACGCTTGTATGTTGGCGGTATCGGTGGTCAATTTG 249
QY 764 GCATCTTTAATATCAAGCAGTGTGCTGGCGCTGCAAGTTCTTGGGTGTGTTCT 823
DB 250 GCTATGAATCACTGAAAAAATTTACGTTTGGCTTGGTGTAGTATGTTGAGCTTGCTG 309
QY 824 ATTGATGCTCCAGATATGTCATGTTCTTGTGTGTCAGTGTGTACCTTCTTCATCGCA 883
DB 310 TTATTCCTGCGAGACAAAGATACATGCAACCGATGATTACTGCTGCAATGCTGCGCA 369
QY 884 TFCGCGCGAGCATGCTTATGCGCTTTACTGCTTGTGTTGCGCAACGCGACATGATCCA 943
DB 370 ATTGCTTTTATCATGCGTGTGTTCTTAACGTTGT-----CTTACGTTTGAAGAT 420
QY 944 GATGCAACCGCTGCTCCAGTGTGCGAGAACGACCAAGCCGAGCAGAACCCCGCA 1003
DB 421 CAACCTAATCCAGAAACAGCAACTGAAAAAAGAGAACTGATAAGATGTTGGCACCCTGTA 480
QY 1004 GAATTTTCAACGATTCACCATCATCCAGCA---CCTTGACCGGTGAAGCTATTGCA 1060
DB 481 AAAACGAATCAAGAGACAAAAATATTTAGCAAGTCCACTTCAAGGTGAAATTTTACCG 540
QY 1061 CTGACGAGCTGAGCGATGCCATGTTTGGCAGCGGAAAGCTTGGCTCGCGGCTGCCATC 1120
DB 541 CTAGAAAAGTACAGACCCCTGTTTGTGCTCAGGTGCTTTAGGAAAAGGTGTTGCAATT 600
QY 1121 GTCCCAACCAAGGCGAGTTAGTTCTTCGGTGTAGTGGAAGATTTGTTGGCANTCCCA 1180
DB 601 GAGCGGACTGAAGGCAAACTGATGACCCCGCAGATGTTGAATCACCACATTTTTCG 660
QY 1181 TCTGCCCATGCTTTGCGAGTTCGACCAAGGCTGAGGATGTTCCAAATGTGGATATCTTG 1240
DB 661 ACAGACATGCTGTTGGCTTGCAGCAACAGAGG-----GGCTTGAATTTA 708
QY 1241 ATGCATTTGTTTTCGACACAGTAACCTCAACGCGCAGCACTTTAAACCCGCTGAGAG 1300
DB 709 ATGCATTTGCGATGATGATGCGTCAATTTAGATGTGTAAGGCTTTTGAATTTACGTGAA 768

QY 1301 CAGGCGCATCAAGTCAAAGCAGGCGAGCTGCTGTGTAATTCGATTTGATGCCATTAA 1360
DB 769 CRAAGTGATCTGTTTAAAAAGAGATTGCTAGTACTTTTGAATTTGCTGCCATTAA 828
QY 1361 GCTGCAGGTTATGAGTAACCAACGCGGATGTTGTTTCGAATTACAA 1407
DB 829 GAAGCTGGTTATCCGCTAGTTTACACCGATTGTCGTAACGAATACGAA 875

RESULT 10
US-08-920-812-17/c
Sequence 17, Application US/08920812
Patent No. 5763188
GENERAL INFORMATION:
APPLICANT: Ohno, Tsuneya
APPLICANT: Matsuhisa, Akio
APPLICANT: Uehara, Hirotsugu
APPLICANT: Eda, Soji
TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/920,812
FILING DATE: 29-AUG-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,577
FILING DATE: 27-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Rin-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 19036/32420
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 3615 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
ORIGINAL SOURCE:
ORGANISM: Escherichia coli
STRAIN: Clinical Isolate EC-24
US-08-920-812-17

Query Match 7.2%; Score 109.6; DB 1; Length 3615;
Best Local Similarity 46.3%; Pred. No. 4.4e-23;
Matches 540; Conservative 0; Mismatches 589; Indels 37; Gaps 4;

QY 312 TCATCCAGTCTGACGTTGCTGCTCACCGGATTCCTTACATTCATCCCAATGSCCAG 371
DB 2495 TCACACATTTGCTATGCTGATGTTATCACACCGTCACCTTTCTGCTGGTGGGCGC 2436
QY 372 CAATCGCGTGGTGGCGATGCTGGCACAGGCTCTACAGGACTTTATGATTTCCGTTG 431
DB 2435 TATCAACTGTGAAGCGAACTGATGCGCGGTTATCTCTGGCTTTATCAGGCGGTT 2376
QY 432 GTCCAGTCCGCGGTTGCTCTTCTGCTGTTGCTCTACTACCAATCGTCATCACTGCTGTC 491

Db 2375 CTGATTTGGGGCGGGTAATGGGGCTTCTGGCAAAATCTTCGTCAATGTCGAACTGC 2316
QY 492 ACAGTCTTCCCGCCAAATGAGCTGAGCTGTTTAACAGGGTGGATCCTTCACTTCG 551
Db 2315 ACTGGGCGCTGGTCCGGTGTGTATCAATTAACCTTCCCGTGTGGGTACGACACCAAGA 2256
QY 552 CAACGGCATCTATG---GCTAATATCGCCAGGGTGGGCAATGTTGGCAGTGTCTTCC 608
Db 2255 TCCGCTGTGTAATCCCGCCATTAATGGCGAGGTCGGGGCGGCTCGGCTTCTCTCT 2196
QY 609 TGGCGAAGAGTGAAGCTCAAGGGCCCTTCAGGTGCTTCAAGTGTCTCCCTGTTCTTG 668
Db 2195 GCGAACCGATGCGCAGAAAAAGTGTGGCGGATCAGCGGCTTACAGAGTCTGTTG 2136
QY 669 GTATTACGGACCTGCGATCTTCGGTGTGAACCTTCGCTCGGTCGGCGGTTCTTCATCG 728
Db 2135 GTATCAGCGAACCGAGGTATATGCGGTCAACCTCGCGCGTAAAGTACCCCTTGTATCG 2076
QY 729 GTATCGGTACCGCAGCTATCGGTGGCGCTTGTGATGCACTCTTTAATCAAGCGAGTTG 788
Db 2075 CCTGTATCAGTGGGCTTGGGGCCACCATATTGGCTACGCGCAACGAAGTCTACT 2016
QY 789 CGTTGGCGCTGCGAGTTTCTTGGGTGTGTTCTATTGATGCTCCAGATATGTCATGT 848
Db 2015 CTTTGGTTTCCAAATATTTTCACTTATGCAAAACCATCCCGTCAACGGGAATGATT 1956
QY 849 TCTTGGTGTGTCAGTGTGTTACCTTCTTCATCGCATTCGGCGCAGCATTCGTTATGCCC 908
Db 1955 TCACCGTCTGGGCCACCGCTTATGCGGTGTCATTCGCCATCGGTTGCGCATTTGTCGTA 1896
QY 909 TTTACTTGGTTCGCGCAAGCGCAGCATTTGATGATGCAAGTCAACCGCTGCTCCAGTGC 968
Db 1895 CGGTGATGCTTCAITTCATCCCGCTAAACGTACGCGCAGCGAGGTGCCCC----- 1844
QY 969 CAGGAACGACCAAAACCGAAGCAGACCGCGAGAAATTTTCAAAAGATTCACCATCA 1028
Db 1843 -----GAAGAGAAACACACAGAGTTATTACACCACTGAGCAGGCGGTA 1798
QY 1029 TCCAGGCACTTTGACCGGTGAGCTATTGCACTGAGCAGCTGAGCATGCAATGTTG 1088
Db 1797 TCTGTTCACCGATGCGGAGAGATTTGTGCTCATTCACGCTGCTGATACAGCTTTG 1738
QY 1089 CCAGCGGAAAGCTTGGCTCGGCGGTGGCATCGTCCCAACCAAGGGCGAGTTAGTTCTC 1148
Db 1737 CAGTGGCTGTTGGGTAAAGTATTGCCATTCGCTCGCTCGTGTGAGTGCCTCTC 1678
QY 1149 CGGTGAGTGAAGATTTGGTGGCATTCCTATCTGGCGCATGCTTTCGAGTTCGACCA 1208
Db 1677 CGGTTGCGGGTCAATTTGCTTGTGTTGCGCACATTACA-----CGCCATTGGCATTG 1624
QY 1209 AGGCTGAGATGTTCCATGTTGGATATCTTGATGCACATTTGTTTCCGACAGTAAACC 1268
Db 1623 AGTCAGATGATGG-----TGTGGAGATCCTGATTCATGTCGGTATCCGACCCGTAAC 1570
QY 1269 TCAACGGCAGCAGCTTAAACCGCTCAAGAGCAGGGCGAATGAAGTCAAGAGCGGGAGC 1328
Db 1569 TGAAGCGCAAAATCTTTCCGCTCAGCTCAACGTGGTCAAGGTCAATACAGCGATC 1510
QY 1329 TCGTGTGAATTCGATATTGATGCAATTAAGGTGCGAGTTATGAGGTAAACAGCCGA 1388
Db 1509 GCGTGAATTTCTTTGATATCCCTGTTATTTCGGAGGCGCGGATTTGATCTGACGACCGG 1450
QY 1389 TTGTTGTTTCGATTAACAGAAAACCGGACCTGTAAACACTTACGTTTGGGCGAATTG 1448
Db 1449 TATTATCATGTAATAGCATGATTTTACGGAGTATATCCCAACGCGCAGCGCAGATAA 1390
QY 1449 AAGCGGAGCCAACTGCTCAACGTC 1474
Db 1389 GCGCAGGTGAACCGCTGTTATCAATC 1364

RESULT 11

US-08-920-827-17/c

; Sequence 17, Application US/08920827
; Patent No. 5770375
; GENERAL INFORMATION:
; APPLICANT: Ohno, Tsuneya
; APPLICANT: Matsuhisa, Akio
; APPLICANT: Uehara, Hirotsugu
; APPLICANT: Eba, Soji
; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/920,827
; FILING DATE: 29-AUG-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,577
; FILING DATE: 27-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 19036/32420
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3615 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli
; STRAIN: Clinical Isolate EC-24
; US-08-920-827-17

Query Match 7.2%; Score 109.6; DB 1; Length 3615;
Best Local Similarity 46.3%; Pred. No. 4.4e-23;
Matches 540; Conservative 0; Mismatches 589; Indels 37; Gaps 4;

QY 312 TCACCTCCAGTGTGACGTTGCTGCACCGGATTCCTTACATTCATGCGCATGGCCAG 371
Db 2495 TCACACCATGTGTATGTCGTATGTTATCACCCGTACCTTTCTGTGGTGGGCGCG 2436
QY 372 CAATCGCTGGTGGCGGATGTGTGGCACACGCTCTACAGGACTTTATGATTTGGTG 431
Db 2435 TATCAACCTGGATAAGCACTGATTCGCGCGGTTATCTCTGCTTATCAGGCGGTT 2376
QY 432 GTCCAGTGGCGGCTGTGCTCTTCGGTCTGTCTACTCACCATTGTCATCATGTCTGC 491
Db 2375 CTGCATTTGCGGGCGCGGTAATGGGCGGCTTTCGCAATCTTCGTCATGTTTCGACTGC 2316
QY 492 ACCAGTCTCTCCGGCAATTGAGCTGGAGCTGTTAAACAGGCTGGATCCTTTCATCTCG 551
Db 2315 ACTGGGCTGTGTCGGTGTGTATCAATAACTTACCGTGTGGCTACGACACCATGA 2256
QY 552 CAACGGCATCTATG---GCTAATATGCCCCAGGGTCGCGCATGTTTGGCAGTGTCTTCC 608
Db 2255 TCCCGCTGTTAATGCCGCCAATTATGGCGAGTCCGGGCGGCTTCGGCGGTCTTCTCTCT 2196
QY 609 TGGCGAAGAGTGAAGAGCTCAAGGGCTTTCGAGGTGCTTCAGGTGTCCTCGCTGTTCTTG 668

Db 2195 GCAGACGATGCGCAAAAAAGTGTGGCGGATCAGCGGGGTTGACGAGTCTGTTG 2136
QY 669 GTATTACGAGCGCTGCGATCTCGGTGTAACCTTCGCTCGCGTGCCTGCTCTTATCG 728
Db 2135 GTATCAGGACAGCGGTATATGGGTCAACTGCGCGTAAGTACCCCTTTGTTATCG 2076
QY 729 GTATCGGTACCGAGCTATCGGTGGCGCTTTGATGCACTCTTTAATATCAAGCGAGTTG 788
Db 2075 CCTGTATCAGTGGGCTTTGGGGCCACCATTTATGGCTACGCGCAAAAGAAAGTCTACT 2016
QY 789 CGTTGGCGCTGCGAGTTCTCGGTGTTGTTCTATTGATGCTCCAGATATGTTATGT 848
Db 2015 CCTTGTGTTGCAAGTATTTTCACTTCATGCAAAACCAATCCCGTCAACGGGAATGATT 1956
QY 849 TCTTGTGTGTGAGTTGTTACCTTCTTCACTGCAATTCGGCGGAGGATTCGTTATGGCC 908
Db 1955 TCACCGTCTGGGCGAGCTTATTTGGGGTGTCAATGCGCATGTTGGCATTTGTGCGTA 1896
QY 909 TTTACTGTTTCCCGCAACGCGAGCATGATCCAGATGCAACCGCTGCTTCCAGTGCCTG 968
Db 1895 CGGTGATGCTTCACTTCACTCCGCTAAACGTCAGCCAGCGAGGTTGCC-----1844
QY 969 CAGGACGACCAAGCCGAGAGAGACCCCGCAATTTCAAGATTCACCATCA 1028
Db 1843 -----GAAGAGAAACCAAGAGTTATTACACCATGAGCGGCGTA 1798
QY 1029 TCCAGGACCTTTGACCGGTGAAGCTATTGCACTGAGCAGCGTCAGCATGCCATCTTTG 1088
Db 1797 TCTGTTACCGATGATCGGAGAGATTTGTGCTGCTATTACGTCGCTGATACACGTTG 1738
QY 1089 CAGCGGAAAGTTGCTCGGGGTTGCCATCGTCCCAACCAAGGGCGAGTTAGTTTCTC 1148
Db 1737 CAGTGGCGCTGTTGGGTAAGGATTTGCCATTCGCTCGGTTGTTGAGTCTCTC 1678
QY 1149 CGGTGAGTGAAGATTTGTTGGCATTCCTCATCTGGCCATGTTTCGAGTTTCGACCA 1208
Db 1677 CGGTGCGGTTGCAATGCTGTTGTTGCGCCACATTA-----CGCCATTGGCATG 1624
QY 1209 AGGCTGAGATGTTTCAATGTTGATCTTGTATGACATGTTTTCGACAGTAACAC 1268
Db 1623 AGTCAGATGATG-----TGTGGAGATCTGATTCATGCTCGGTATCGACCGTAAAC 1570
QY 1269 TCAACGCGACGACTTTAACCGCTGAAGAGCAGCGCGATGAAGTCAAGCAGGGAGC 1328
Db 1569 TGGACGGCAAAATCTTTTCGGCTCAGTCAACGTTGCTGACAGGTCAATACAGCGATC 1510
QY 1329 TGCTGTGTGAATTCGATATTGATGCTCAATTAAGCTGCGAGTTATGAGGTAAACAGCGGA 1388
Db 1509 GGCTGATTTCTTTTGAATTCCTGTTATTCGAGGCGCGGATTTGATCTGACGAGCGCG 1450
QY 1389 TTGTTGTTTTCGAATTACAGAAACCGGACCTGTAACACTTACGTTTGGGCGAAATTG 1448
Db 1449 TATTATCAGTAATAGCATGATTTTACGAGATATTACCCACGCGACGCGCGCATAA 1390
QY 1449 AAGCGGAGCCAACTGCTCAACGTC 1474
Db 1389 GCGAGGTGAACCGTGTATTCATC 1364

RESULT 12
US-08-921-177-17/c
Sequence 17, Application US/08921177
Patent No. 5798211
GENERAL INFORMATION:
APPLICANT: Ohno, Tsuneya
APPLICANT: Matsuhisa, Akio
APPLICANT: Uehara, Hirotsugu
APPLICANT: Eda, Soji
TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun

STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: 29-AUG-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA: US 08/362,577
FILING DATE: 27-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Rin-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 19036/32420
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 3615 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
ORGANISM: Escherichia coli
STRAIN: Clinical Isolate EC-24
US-08-921-177-17

Query Match 7.2%; Score 109.6; DB 1; Length 3615;
Best Local Similarity 46.3%; Pred. No. 4.4e-23;
Matches 540; Conservative 0; Mismatches 589; Indels 37; Gaps 4;
QY 312 TCATCCAGTGTGAGTGTGCTCACCGAATTCCTTACATTCATCGCATTCGCCAG 371
Db 2495 TCACACCATGCTATGCTGATGTTATCACCCGTACCTTCTGCTGGTGGGCCGC 2436
QY 372 CAATCGCTGGTGGCGAGTGTGGCACACGGTCTACAGGACCTTATGATTTCCGTG 431
Db 2435 TATCAACTGGATAAGCAACTGATTGCGCGGTTATCTCTGGCTTATCAGGGGTT 2376
QY 432 GTCCAGTGGCGGCTGCTCTTCGCTGCTGCTTACTCACCATGCTCATCTGCTG 491
Db 2375 CTGCAATTTGGCGCGCGGTAAATGGCGCGGCTTCTGGCAATCTTCGTCATGTT 2316
QY 492 ACCAGTCTTCCCGCAATTCAGCTGGAGCTGTTTAAACAGGGTGGATCTTTCATCT 551
Db 2315 ACTGGGCGCTGCTGGCGGTGTATCAATACTTCCGCTGCTGGCTACGACATGA 2256
QY 552 CAACGGCATCTATG--GCTAATATCGCCAGGTGGCGCATGTTTGGCAGTGTCTCC 608
Db 2255 TCCCGCTGTAAATGCCGCCATTATGGCGAGGTGGCGCGGCTCGCGGCTTCTCT 2196
QY 609 TGGCGAAGTGAAGCTCAAGGCGCTTCAAGGCTTCAAGGCTTCAAGTGTCTGTTG 668
Db 2195 GCGAACGATGCGCAGAAAGTGGTGGCGGATCAGCGCGGTGACGAGTCTGTTG 2136
QY 669 GTATTACGAGCTGCGCATCTTCGCTGTGAACCTTCGCTGCGCTGGCGCTTCTTC 728
Db 2135 GTATCAGGACAGCGGTATATGGGTCAACCTTCCGCGTAAGTACCCCTTTGTTAT 2076
QY 729 GTATCGGTACCGAGCTATCGGTGGCGCTTTGATGCACTCTTTAATATCAAGGAG 788
Db 2075 CCTGTATCAGTGGGCTTTGGGGGCCACCATTTATGGCTACGCGCAAGAGTCTACT 2016

Db	1677	CGGTTGGGTCGAATTCCTGTTTCGCCCAATTACA-----CGCCATTGGCAATTG	1624
QY	1209	AGGCTGAGGATGGTTCCAAATGTGGATATCTTGATGACATGTTGGTTTCGACACAGTAAACC	1268
Db	1623	AGTCAGATGATGG-----TGTGGAGATCCTGATTCATGTCGGTATCGACACACCGTAAAC	1570
QY	1269	TCAACGGCACGCACTTTAAACCCGCTGAAGAAGCAGGCGGATGAAGTCAAAAGCAGGGAGC	1328
Db	1569	TGGACGGCAAAATCTTTTCGGCTCACGTCAACGCTGGGTGAACAAGTCAATACACGGCGATC	1510
QY	1329	TGCTGTGATTCGATTTGATTCGCCATTAAGCTGCAGGTTATGAGTAAACCAACGCGGA	1388
Db	1509	GGCTGATTTCTTTTGATATCCCTGCTATTCGAGGCGCGGATTTGATCTGACGACGCGG	1450
QY	1389	TTGTTGTTTCGAATTACAAGAAAACCGGACCTGTAACACACTTACGGTTTGGCGAAATTTG	1448
Db	1449	TATTAATCAGTAATAGCATGATTTTACGGACGTTATTACCCACGCGCACGCGCAGATAA	1390
QY	1449	AAGCGGAGGCAACTGCTCAACGTC	1474
Db	1389	GCGCAGGTGAACCGCTGTATTCCATC	1364

RESULT 15
US-09-543-681A-2847
; Sequence 2847, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 2847
; LENGTH: 1905
; TYPE: DNA
; ORGANISM: Proteus mirabilis
US-09-543-681A-2847

Query Match	7.2%;	Score 109.4;	DB 4;	Length 1905;
Best Local Similarity	48.3%;	Pred. No. 3.4e-23;	Mismatches 421;	Indels 18; Gaps 3;
Matches	410;	Conservative		

QY	567	CTAATATCGCCAGGTCGGCATGTTTGGCAGTGTCTTCTCGGGAAGAGTGAAGAAGC	626
Db	1004	CCAAACATGTCCTCAGGCTGCTGCTTTCTTTAGCTGTTTCTGTAGTACTAAATAAACCCT	1063
QY	627	TCAAGGGCCCTTCAGGTCGTTTCAGGTGTCTCCGCTGTTCT---TGCTATTACGGACCTG	683
Db	1064	TAAACAATTAGCTTTTCTGCCAGTATTACCGCTTTTCTTTGGTGTATAAATGAACCCG	1123
QY	684	CGATCTTCGGTGTGAACCTTCGCTGCGCTGGCGCTTCTTCATCGTATCGGTACCGCAG	743
Db	1124	CTATGTATGGGTGAACCTCAAACTGAAAAAACCAATGACGCTGTGTATGATTGGTGGTG	1183
QY	744	CTATCGGTGGCGCTTTGATTGCACTCTTAAATCAAGCAGTTCGCTTGGCGCTGCAG	803
Db	1184	CTAATGCAGGATTTTGGTGGTATTGTTAAGTTAAAGGTTTGTCTATGTCACCTCCGG	1243
QY	804	GTTTCTTGGGTGTTGTTTCTAATGATGCTCCAGATGTTGTCATGTTCTTGGTGTGCGAG	863
Db	1244	GGTTAATAAG---TTTACCAATGTGGATATCTGATACAGATAACCAAGTAGTAATGCA	1300
QY	864	TTGTTACCTTCTTCATCGCAATTCGGCGCAGGATTTGCTTATGGCTTTACTTGGTTCGCC	923
Db	1301	TAATCACCTTATTGATTGGCAGTGTAGCCACTTTTATTGCAACGCTTATTATTGGTTTG	1360
QY	924	GCAACGGCAGCAATTGATCCAGATGCAACCGCTGCTCCAGTGCCTGCAGGAACGACCAAG	983

Db	1361	ATGATCCTACTGATGACCCCAATCCGTGACGAAGAAAAATAATAACAAGCTGCATCAA	1420
QY	984	CCGAAGCAGGAAGCACCGCGAGAAATTTTCAAACGATTCCACCATCATCCAGGCACCTTTGA	1043
Db	1421	CCAACAAAAAACCAAAATAGCGAATAGTAATTTACCTGTGGTTAATTTTCACTCTGC	1480
QY	1044	CCGCTGAAGCTATTGCACTGAGCAGCGTCAAGCATGTTTTCAGCGGAAAGCTTG	1103
Db	1481	AAGTAAACCGTTGCACTTTCAGAGGTTAATGATGAACATTTGCTCTGGCATTATGG	1540
QY	1104	GCTCGGCGTTGCACTGCTCCCAACCAAGGGCGAGTTAGTTTCTCCGGTGAAGTGAAGA	1163
Db	1541	GACCTGGCATGGCAATTTATCCCAACACAGGAAAGTGATTCACCTGCGAGTGGCGTTG	1600
QY	1164	TTGTGGTGGCATTCCTCATCTGGCCATCTTTTCGCAGTTCGCCAACAGGCTGAGGATGTT	1223
Db	1601	TTGATATTACCTTTTCTTCTGGTCATGCGATTGGCTTAAACACTG-----GTTA	1648
QY	1224	CCATGTGGATATCTTGATGCACATTTGTTTCGACACAGTAACCTCAACGCGCAGCACT	1283
Db	1649	ACATATTGAAATGCTGATCCATGCGGATTTGATACCGTTAACTTAGCGGCGCAGCACT	1708
QY	1284	TTAACCCGCTGAAGACAGGCGCATGAGTCAAGCAGGGGAGCTGCTGTGAATTG	1343
Db	1709	TCATATGCTGTGCTGTTAAAGGTCAGAAAGGTGACAAAGGTTGATACGCTAATTGAATTG	1768
QY	1344	ATATTGATGCCATTAAAGGCTGCAGTTATGAGGTAAACCAACCGCGATTGTTGTTTGAATT	1403
Db	1769	ATTAGATGCGATTAATAGCCGAGCTATGATCCACAAACCGATGATCATCATTAATG	1828
QY	1404	ACAAGAAAA	1412
Db	1829	ATGATCATA	1837

Search completed: October 30, 2004, 17:07:49
Job time : 166 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 30, 2004, 15:50:57 ; Search time 744 Seconds
(without alignments)
10523.989 Million cell updates/sec

Title: US-09-604-231-1

Perfect score: 1527

Sequence: 1 cccatggatctgcgcgtt.....gttgaaacctgagtgttcg 1527

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3413475 seqs, 2563800928 residues

Total number of hits satisfying chosen parameters: 6826950

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Length	ID	Description
1	1527	100.0	1527	16	US-10-450-055-9
2	1515	99.2	3309400	9	US-09-738-626-1
3	1492	97.7	1983	9	US-09-738-626-2904
4	177.8	11.6	975	9	US-09-974-300-829
5	151.2	9.9	5840	9	US-09-070-927A-31
6	137.2	9.0	1380	9	US-09-974-300-685
7	135.4	8.9	3037	16	US-10-398-221-3599
8	134.6	8.8	684707	16	US-10-398-221-9
9	134.6	8.8	3011208	16	US-10-398-221-2058
10	126.6	8.3	2127	17	US-10-474-776-160
11	126.6	8.3	3895	8	US-08-961-527-201
12	126.6	8.3	3895	16	US-10-158-844-201

13	113.4	7.4	1884	9	US-09-815-242-9369	Sequence 9369, App
14	107.8	7.1	1185	17	US-10-474-776-152	Sequence 152, App
15	107.8	7.1	9769	8	US-08-961-527-30	Sequence 30, Appl
16	107.8	7.1	9769	16	US-10-158-844-30	Sequence 822, App
17	105.6	6.9	1014	9	US-09-974-300-822	Sequence 822, App
18	99	6.5	1098	9	US-09-974-300-734	Sequence 734, App
19	98	6.4	489	9	US-09-974-300-857	Sequence 857, App
20	93.6	6.1	1854	16	US-10-398-221-712	Sequence 712, App
21	93.6	6.1	1854	16	US-10-398-221-2122	Sequence 2122, App
22	93.6	6.1	1854	16	US-10-282-122A-24691	Sequence 24691, A
23	87.8	5.7	1137	16	US-10-282-122A-34025	Sequence 34025, A
24	86.2	5.6	30246	8	US-08-781-986A-56	Sequence 56, Appl
25	86.2	5.6	30246	16	US-10-329-624-56	Sequence 56, Appl
26	81.6	5.3	1854	16	US-10-398-221-1040	Sequence 1040, App
27	81.6	5.3	1854	16	US-10-398-221-2857	Sequence 2857, App
28	78.8	5.2	12278	16	US-10-398-221-3886	Sequence 3886, App
29	77.4	5.1	444	9	US-09-974-300-877	Sequence 877, App
30	76.8	5.0	1368	9	US-09-974-300-893	Sequence 893, App
31	76.4	5.0	465	8	US-08-781-986A-1317	Sequence 1317, App
32	76.4	5.0	465	16	US-10-329-624-1317	Sequence 1317, App
33	74.2	4.9	474	9	US-09-070-927A-639	Sequence 639, App
34	73.2	4.8	1357	16	US-10-282-122A-21444	Sequence 21444, A
35	72.4	4.7	1953	16	US-10-282-122A-38846	Sequence 38846, A
36	72.4	4.7	1956	16	US-10-282-122A-36728	Sequence 36728, A
37	70.4	4.6	9797	9	US-09-070-927A-550	Sequence 550, App
38	69.8	4.6	1353	9	US-09-974-300-799	Sequence 799, App
39	69.4	4.5	429	9	US-09-974-300-5182	Sequence 5182, App
40	69.2	4.5	1953	16	US-10-282-122A-39556	Sequence 39556, A
41	68.2	4.4	348	9	US-09-974-300-5313	Sequence 5313, App
42	67.2	4.4	2028	16	US-10-282-122A-16373	Sequence 16373, A
43	67	4.4	2181	17	US-10-474-776-65	Sequence 65, Appl
44	67	4.4	8494	8	US-08-961-527-163	Sequence 163, App
45	67	4.4	8494	16	US-10-158-844-163	Sequence 163, App

ALIGNMENTS

RESULT 1
US-10-450-055-9
; Sequence 9, Application US/10450055
; Publication No. US20040043953A1
; GENERAL INFORMATION:
; APPLICANT: BASF Aktiengesellschaft
; TITLE OF INVENTION: No. US20040043953A1el genes of Corynebacterium
; FILE REFERENCE: 936 2000
; CURRENT APPLICATION NUMBER: US/10/450,055
; CURRENT FILING DATE: 2003-06-10
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Vers. 2.0
; SEQ ID NO 9
; LENGTH: 1527
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (101)..(1504)
; OTHER INFORMATION: RXS00315
US-10-450-055-9

Query Match	100.0%	Score 1527;	DB 16;	Length 1527;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 1527;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	CTCATGGCATCTGCGCGTTCGCGTTCCTTCCAGTGTGTGTTTCCACGCAACCAAG	60	
Db	1	CTCATGGCATCTGCGCGTTCGCGTTCCTTCCAGTGTGTGTTTCCACGCAACCAAG	60	
Qy	61	CGTTTCGGCGCATAGTTCCTGGGCGCCCGCTATTGGTATGGCGATGGTGTCCCGAG	120	
Db	61	CGTTTCGGCGCAATAGTTCCTGGGCGCCCGCTATTGGTATGGCGATGGTGTCCCGAG	120	
Qy	121	CTTGGTGAACGGCTACGACGTGGCGGCCACCATGGCTCGCGCGCAATGCCAATGTGGTC	180	

Db 121 CTTGGTGAACGGCTACGAGTGGCGGCCACATGGCTGCGGCGAATGCAATGTGTC 180
Qy 181 CTTGTTTGGTTAGATTGCTCCCAAGCCGTTACAGGGCAACCGTGTCTCTGCTGTGTT 240
Db 181 CTTGTTTGGTTAGATTGCTCCCAAGCCGTTACAGGGCAACCGTGTCTCTGCTGTGTT 240
Qy 241 GGTTCCTTGGATTCTGCAACGATCGAAGATTCCTGCAAGGCACTCAAGGCACTGC 300
Db 241 GGTTCCTTGGATTCTGCAACGATCGAAGATTCCTGCAAGGCACTCAAGGCACTGC 300
Qy 301 AGACTTCCTGATCACTCCAGTGTGAGCTGTCTCTCTCACCGGATTCCTTCAATTCATCGC 360
Db 301 AGACTTCCTGATCACTCCAGTGTGAGCTGTCTCTCTCACCGGATTCCTTCAATTCATCGC 360
Qy 361 CATTGGCCCAAGCAATCGCTGGGTGGCGATGTCTGTGCAACGCTCAAGGCACTTTA 420
Db 361 CATTGGCCCAAGCAATCGCTGGGTGGCGATGTCTGTGCAACGCTCAAGGCACTTTA 420
Qy 421 TGATTTCGTTGGTCCAGTCCGCGGTCTGCTCTTCGTTCTGCTTACTCAACATCGTCAT 480
Db 421 TGATTTCGTTGGTCCAGTCCGCGGTCTGCTCTTCGTTCTGCTTACTCAACATCGTCAT 480
Qy 481 CACTGGTCTGCACAGTCTCTCCGCGCAATTTGAGCTGGAGCTGTTTAAACAGGCTGGATC 540
Db 481 CACTGGTCTGCACAGTCTCTCCGCGCAATTTGAGCTGGAGCTGTTTAAACAGGCTGGATC 540
Qy 541 CTTTCATCTTCGCAACGGCATCTATGGCTAATATCGCCAGGTCGGCATGTTTGGCACT 600
Db 541 CTTTCATCTTCGCAACGGCATCTATGGCTAATATCGCCAGGTCGGCATGTTTGGCACT 600
Qy 601 GTTCTCTCGGCAAGAGTGAAGCTCAAGGCTTCAGGCTTCAGTGTCTTCGCTCCGC 660
Db 601 GTTCTCTCGGCAAGAGTGAAGCTCAAGGCTTCAGGCTTCAGTGTCTTCGCTCCGC 660
Qy 661 TGTTCCTGTTATACGAGCTTCGCAATCTTCGTTGTGAACCTTCGCTTCGCTGGCCGTT 720
Db 661 TGTTCCTGTTATACGAGCTTCGCAATCTTCGTTGTGAACCTTCGCTTCGCTGGCCGTT 720
Qy 721 CTTTCATCGTTACGTTACGAGCTATCGGTGGCGCTTTGATGCACTCTTTAATATCAA 780
Db 721 CTTTCATCGTTACGTTACGAGCTATCGGTGGCGCTTTGATGCACTCTTTAATATCAA 780
Qy 781 GGCAGTTGGTTGGCGCTTCAGGCTTCGTTGGTGTGTTTCTATTGATGCTCCAGATAT 840
Db 781 GGCAGTTGGTTGGCGCTTCAGGCTTCGTTGGTGTGTTTCTATTGATGCTCCAGATAT 840
Qy 841 GGTCAATGTTCTGTTGTGCAAGTGTGTTACCTTCCTTCATCGCATTCGGCGAGCGATTGC 900
Db 841 GGTCAATGTTCTGTTGTGCAAGTGTGTTACCTTCCTTCATCGCATTCGGCGAGCGATTGC 900
Qy 901 TTATGCGCTTTACTTGGTTTCGCGCAACGCGAGCTATGATCGATGCAACGCTGCTCC 960
Db 901 TTATGCGCTTTACTTGGTTTCGCGCAACGCGAGCTATGATCGATGCAACGCTGCTCC 960
Qy 961 AGTGGCTGCAGAAACGACCAAGCCGAGCAAGCAGCAGCAGATTTTCAAAAGATTC 1020
Db 961 AGTGGCTGCAGAAACGACCAAGCCGAGCAAGCAGCAGCAGATTTTCAAAAGATTC 1020
Qy 1021 CACCATCAATCCAGGCACTTTGACCGGTTGAAGCTATTCACCTGAGCAGCTCAGCGATGC 1080
Db 1021 CACCATCAATCCAGGCACTTTGACCGGTTGAAGCTATTCACCTGAGCAGCTCAGCGATGC 1080
Qy 1081 CATGTTTGCAGGGAAGCTTGGCTCGGGGTGTCATCGTCCCAACCAAGGGGCACTT 1140
Db 1081 CATGTTTGCAGGGAAGCTTGGCTCGGGGTGTCATCGTCCCAACCAAGGGGCACTT 1140
Qy 1141 AGTTTCTCGGTGAGTGAAGATTCGTGTGGCAATTCCTGCTGGCCATGCTTTTCCAGT 1200
Db 1141 AGTTTCTCGGTGAGTGAAGATTCGTGTGGCAATTCCTGCTGGCCATGCTTTTCCAGT 1200
Qy 1201 TCGCCAAAGGCTGAGGATGGTTTCCAAATGTGGATATCTTTGATGCAATTTGGTTTCGAC 1260

Db 1201 TCGCACCAAGGCTGAGGATGGTTCCAAATGTGGATATCTTGTATGCAATTTGGTTTCGACAC 1260
Qy 1261 AGTAAACCTCAACGGCAACGCTTTTAAACCCGCTGAAGAACGAGCGGCGATGAAGTCAAAGC 1320
Db 1261 AGTAAACCTCAACGGCAACGCTTTTAAACCCGCTGAAGAACGAGCGGCGATGAAGTCAAAGC 1320
Qy 1321 AGGGAGCTCTCTGTGTGAATTCGATATTTGATGCGCATTTAAGGCTGCGAGTTATGAGGTAAC 1380
Db 1321 AGGGAGCTCTCTGTGTGAATTCGATATTTGATGCGCATTTAAGGCTGCGAGTTATGAGGTAAC 1380
Qy 1381 CAGCCGATTTGTTTTCGAAATTCAGAAACCGGACCTGTAAACACTTACGTTTGGG 1440
Db 1381 CAGCCGATTTGTTTTCGAAATTCAGAAACCGGACCTGTAAACACTTACGTTTGGG 1440
Qy 1441 CGAAATTTGAAGCGGAGCCCAACCTGCTCAACGTCGCAAAAGAAAGCGGTGCCAGCAAC 1500
Db 1441 CGAAATTTGAAGCGGAGCCCAACCTGCTCAACGTCGCAAAAGAAAGCGGTGCCAGCAAC 1500
Qy 1501 ACCATAAGTTGAAACCTTTGAGTGTTCG 1527
Db 1501 ACCATAAGTTGAAACCTTTGAGTGTTCG 1527

RESULT 2
US-09-738-626-1/c
; Sequence 1, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-16
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 1
; LENGTH: 3309400
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-1

Query Match 99.2%; Score 1515; DB 9; Length 3309400;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1526; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 CTTATGGCATCTGCGCGTTCGCTTCTTCCAGTGTGTTTCCCGCAACCAAG 60
Db 2811326 CTTATGGCATCTGCGCGTTCGCTTCTTCCAGTGTGTTTCCCGCAACCAAG 2811267
Qy 61 CTTTCGCGGCAATGAGTTCCTGGGCGCG-GTATTGTAAGCGAUGTGTCCCGAG 120
Db 2811266 CTTTCGCGGCAATGAGTTCCTGGGCGCG-GTATTGTAAGCGAUGTGTCCCGAG 2811208
Qy 121 CTTTGTGAACGGCTACGACGTGGCGCCACCATGCTCGGCGCAAAATGCCAATGTGTC 180
Db 2811207 CTTTGTGAACGGCTACGACGTGGCGCCACCATGCTCGGCGCAAAATGCCAATGTGTC 2811148
Qy 181 CTTGTTGGTTTAGATGTTGGCCCAAGCCGTTACCAAGGCGACCGTCTTCTTGTGCTGGT 240

Db 466 GAACAATTCGGCTACTGGCGGCACTGTTATACGAGGGCTTTACGCCGCTCTCGTCATT 525
QY 482 ACTGCTCTGCACAGCTCTTCCGCCAAATGAGCTGGAGCTGTT-----AACAGGGT 535
Db 526 ACCGGAATGATCATACATCTCTTGGCTTGAGCTTCAGCTGATCGGCTCAAGCTCGGC 585
QY 536 GATCTCTTCATCTCCGACCGCATCTATGGCTPAATATCGCCAGGGTGGCGCATGTTG 595
Db 586 GGAACATATTTTATGGCGGATGCTCGCGCTGTCATATATGACAAAGGTTGAGCGCGCT 645
QY 596 GCAGCTCTCTCTGCGGAGAGTGAAGAGCTCAAGGGCTTCAGGTGCTTCAGGTGTC 655
Db 646 GCATGATGTTTATGTCGAAGATGAGAGCGAAGAGGCTGTCCTGACATCCGAAAT 705
QY 656 TCGCTGTTCTTGGTATTACGGAGCCTGCGATCTTCGGGTGTGAACCTTCGCCCTGCGCTG 715
Db 706 TCAGCTTATCTCGGAATTACCGAACCAGCATTTGTCGGAGTGAATCTCAGATACAAAT 765
QY 716 CCGTCTTCATCGTATCGGTACCGCAGCTATCGGTGGCGCTTTGATTCACCTCTTTAAT 775
Db 766 CCGTTGCTATCGCATGATCATGATTCGGGATTCGCGGTATGTTTCTATTTCTCGCAAGG 825
QY 776 ATCAAGCGAGTTCGTTGGCGCTGCGAGCTTCTTGGGTGTTGTTCTATT 826
Db 826 GTTTGGCAAGCTCTGTCGGCGTGGCGGCGTACCTGGAAATTTCTCTATT 876

RESULT 5

US-09-070-927A-31
; Sequence 31, Application US/09070927A
; Patent No. US20020120116A1
; GENERAL INFORMATION:
; APPLICANT: Charles A. Kunsch
; Patrick J. Dillon
; Steven Barash
; TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 982
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/070,927A
; FILING DATE: 04-May-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/046,655
; FILING DATE: 1997-05-16
; APPLICATION NUMBER: 60/044,031
; FILING DATE: 1997-05-06
; APPLICATION NUMBER: 60/066,009
; FILING DATE: 1997-11-14
; ATTORNEY/AGENT INFORMATION:
; NAME: Kenley K. Hoover
; REGISTRATION NUMBER: 40,302
; REFERENCE/DOCKET NUMBER: PB369
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5840 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear

; SEQUENCE DESCRIPTION: SEQ ID NO: 31:

US-09-070-927A-31

Query Match 9.9%; Score 151.2; DB 9; Length 5840;
Best Local Similarity 48.5%; Pred. No. 8.5e-36;
Matches 525; Conservative 0; Mismatches 533; Indels 24; Gaps 3;
QY 329 TTGCTGCTCACCGGATTCCTTACATTCATCGCCATGGCCAGCAATGCGCTGGTGGG 388
Db 1850 TTCTTATATGCGACCTGTTACCTTCTTAGCTATTGGTCTTATCGGGAGCGTTCATCGGC 1909
QY 389 GATGCTCTGGCACACGGTCTACAGGACCTTATGATTTCCGGTGTTCAGTTCGCGGCTG 448
Db 1910 GATGCTTGGGACAGGATACACGCAATTTACGCTTTTAGTCCAAATATCGCAGGGTTA 1969
QY 449 CTCCTGCTGCTGCTCTACTCAACCAATCGTCACTGCTGTCACCAAGTCTTCCGCGCA 508
Db 1970 TTAATGGGTTCTGTTGGCAAGTCTTGGTAAATGTTGGTATGCAATGGGCTTTGTACCA 2029
QY 509 ATTGAGCTGAGCTGTTTAAACAGGGTGGATCCTTCATCTTCGCAACGGCACTATGGCT 568
Db 2030 ATTATGATGTTTAACTTAAACAAGGTGGCATACGATGGTACCGATGTTTATACAGCC 2089
QY 569 AATATCGCCAGGGTGGCGCATGTTTGGCAGTGTCTTCTCGCGAAGAGTGAAGAGCTC 628
Db 2090 GTTATGACAAAGCGGGCTGCTTTAGCTGCTTTTCTTAACAAAATATGGAACATA 2149
QY 629 AAGGGCTTGGAGGTGCTTCAAGGTCTCCGCTGTTCTTGGTATTACGAGGCTGCGATC 688
Db 2150 AAAGGTTTGGCTTTGCTTCAAGTATTACGACTATTTTGGAAATTAATGAACTGTA 2209
QY 689 TTCGGGTGAACCTTCGCTGCGCTGGCGCTTCTTCATCGGTATCGGTACCGCAGCTATC 748
Db 2210 FATGGCTGACTTTTACCATTGAAAAAACCATTATTTCAGCTGTGTTATGCTGGCGTATC 2269
QY 749 GGTGGCGCTTTGATTCGACTCTTTTAATATCAAGGAGTTCGCTTGGCGCTGCGAGTTTC 808
Db 2270 GGTGGTGCATTTGTCGTATGATCAAGTGAAGAACTTTACGTTTGGCTTGGTGTATG 2329
QY 809 TTGGGTGTTGTTTCTTATGATGCTCCAGATGCTGTCATGTTCTTGGTGTGTCAGTGT 868
Db 2330 TTGAGCTTGGCTGGCTTTTATCTCGCAGAGACAAAAGATCTGCACCGATGATTAATGCT 2389
QY 869 ACCTTCTTTCATCGCATTCGGCGCAGGATTCGCTTATGGCTTTACTTGGTTCGCGCAAC 928
Db 2390 GCAATGGTGGCGGAATTCGCTTTATCATGCTTGTGCTTAAACGTTTGTCTTACGTT-- 2447
QY 929 GGCAGCATTTGATCCAGATCCCAACCGCTGCTCCAGTGCCTGCGAGCAACGACCAAGCGAA 988
Db 2448 -----TTGAAGATCAACCTAATCCGAACAGCAACTGAAAAACAGAACTGATAAG 2500
QY 989 GCAGAAGCACCGCGAGAAATTTTCAACGATTCACATCCATCCAGGCA---CCTTTGACC 1045
Db 2501 ATGTGGCACCTGTAAAAACCAATCAAGAGACAAAAATTTTATAGCAAGTCCACTCAA 2560
QY 1046 GGTGAAGCTATTGCACTGAGCAGGCTCAGCGATGTCATGTTTGCAGCGGAAAGTTGGC 1105
Db 2561 GGTGAATTTTACCGCTAGAAAAAGTACAACCTGTTTTCCTTCAGGTGCTTTAGGA 2620
QY 1106 TCGGCGTGGCATGCTGCCCAACCAAGGGGAGTTAGTTTCTCCGGTCACTGCGAAAGATT 1165
Db 2621 AAAGGTGTTGCAATTTAGCCGACTGAAGCAAACTGATATGACCCGCGAGATGGTGAATC 2680
QY 1166 GTGTGGCAATTCCTCATCTGGCCATGCTTTTCAGTTCGACCAAGGCTGAGATGTTTCC 1225
Db 2681 ACCACATTTATTCGACAGGACATGCTTGGCTTGAACCAACAGAGGGC----- 2731
QY 1226 AATGTGGATATTTGATGCAATTTGTTTCGACACAGTAACCTCAACGGCAGCACATTT 1285
Db 2732 ---GTTGAATTTAATGTCATATTTGGCATGATGATCGGTGGAATTAGATGCTAAAGGCTTT 2788
QY 1286 AACCCGCTGAAGAAGCAGGCGGATGAAGTCAAGCAGGGAGCTCTCTGTGAATTCGAT 1345

QY 853 GGTGTGTCAGTGTGTTACCTTCTTCATCGCAATCGGCGCAGCATGCTTATGGCCCTTTA 912
 Db 1218 CGAATTCGTGGGTAGTTATCGCAATCGTTATCAGCTTCTATCTTGGTTCTATTTAAC 1277
 QY 913 CTGTGTGTCGGCGCAACGGCAGCAATGATCAGATGCAACCGCTGCTCCAGTGCCTGCAGG 972
 Db 1278 TTATGTA-----GTAGGTTTCAAGATCCAGCTGATGTAGTAGTAGAACA 1322
 QY 973 AACGACCAAGCCGAAGCAGACACCGCAGAAATTTTCAAGATTTCCACCACATCTCCA 1032
 Db 1323 ATCTAACACAGTAGAGGCGAAGC-----TTTAAATCGAAGTGAACAATTC 1370
 QY 1033 GGCACCTTTGACCGGTGAAGCTATTCGACTGAGCAGCGTCAGCGATGCCATGTTTGCAG 1092
 Db 1371 AGCTCCAGTAGTTGGCGAAATCGTAACTTTAGCAGACGTTAAAGAGCGAACAATTTTCATC 1430
 QY 1093 CGGAAAGCTTGGCTCGGCGGTGCGCATCGCTCCCAACCAAGGGCGAGTTAGTTCTCGGT 1152
 Db 1431 CGGTGCACCTTGGAAAGGTGTTCGAATCATCCCAACTGTTGGACGTGTAGTAGTCCAGC 1490
 QY 1153 GAGTGGAAAGATGTGGTGGCATTCCTCATCTGCCCATGCTTTCGCAGTTTCGCACCAAGGC 1212
 Db 1491 AGCAGGAACAGTAACACTATCTTCCCACTGGTCAATCGGTATTAACAATAAG- 1549
 QY 1213 TGAGGATGGTCCAAATGTGGATATCTTGATGCAATTTGTTTCGACACAGTAACCTCAA 1272
 Db 1550 -----ACGGTGCAGAAAGTTTAAATCCACATCGGTATGACACCCGTTCAACTTGA 1598
 QY 1273 CGGCGCAGCTTTAAACCCCTGAGAGCAGGCGGATGAAGTCAAGAGCGGGAGCTGCT 1332
 Db 1599 AGGTAAATTTCTTACAGCACAAGTGAACCAAGCGACGTAATCGAAAAAGGTCAATTAAT 1658
 QY 1333 GTGTGAATTCGATATGATGTCATTAAGCTGCGAGTTATGAGTAAACCAAGCGGATTCG 1392
 Db 1659 AACTGAATTTGATTTGAAGGCACTAAGAGCTGCTGGAATGACGTTTACAACTCCAGTAGT 1718
 QY 1393 TGTTCGAATTAAGAAACCGGACCTGTAAACACT 1429
 Db 1719 AGTAACAAACTCCAATCAATATTAGACGTAATGATT 1755

RESULT 8
 US-10-398-221-9/c
 ; Sequence 9, Application US/10398221
 ; Publication No. US2004018514A1
 ; GENERAL INFORMATION:
 ; APPLICANT: KUNST, Frederik
 ; APPLICANT: GLASER, Philippe
 ; TITLE OF INVENTION: Listeria innocua, genome and applications
 ; FILE REFERENCE: 344 702 - US
 ; CURRENT APPLICATION NUMBER: US/10/398,221
 ; CURRENT FILING DATE: 2003-03-27
 ; PRIOR FILING DATE: PCT/FR 01/03 061
 ; PRIOR FILING DATE: 2001-10-04
 ; PRIOR APPLICATION NUMBER: FR 00/12 697
 ; PRIOR FILING DATE: 2000-10-04
 ; NUMBER OF SEQ ID NOS: 4025
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 9
 ; LENGTH: 684707
 ; TYPE: DNA
 ; ORGANISM: Listeria innocua
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1)-(end)
 ; OTHER INFORMATION: n can be any nucleotide: a, g, c or t/u

Query Match 8.8%; Score 134.6; DB 16; Length 684707;
 Best Local Similarity 48.2%; Pred No. 1.3e-29;
 Matches 537; Conservative 0; Mismatches 534; Indels 42; Gaps 4;
 US-10-398-221-9

QY 348 TTACATTCATCGCAATTCGGCCAGCAATCGCTGGGTGGGATGTCTGGCACACCGCTC 407
 Db 240932 TTACTTTCCTTGTAAATCGGACCAATCGCAATCTGGGCTGTGTCAATTAATCTTGGAGCAGGTA 240873
 QY 408 TACAGGACCTTTATGATTTTCGGTGGTCCAGTCGCGGTCTGTCTCTTCGGTCTGGTCTACT 467
 Db 240872 CAATTTGGGTTTATAACTTAAGCCAGTGTGTCTGTTAACTTCTAGGTGGTTCCTGGC 240813
 QY 468 CACCAATCGTCATCAGTGTCTGACAGTCCCTTCGCGCAATTTGAGCTGGAGCTGTTTA 527
 Db 240812 AAGTATTCGTTATCTTTGGACTTCACTGGGCTCTTATTCAGTGGCAATCAACAACCTAA 240753
 QY 528 ACCAGGCTGATCCTTCA---TCTTCGCAACGGCATCTATGGCTATATATCGCCAGGGTG 584
 Db 240752 CAACGCTTGAATCAGATCCAGTACTTGGCAATGATGTTCGGTGTCTCTTTTGGCCAAATCG 240693
 QY 585 CGGCATGTTTGGCAGTGTCTTCTTCGCGCAAGAGTGAAGAGCTCAAGGGGCTTTCAGGTTG 644
 Db 240692 GTGCAGTATTAGCAGTATCTTCAAACTAGAAACAAAAAAATCAATCTCTTAGTATTC 240633
 QY 645 CTTTCAGGTGTCTCGCTGTCTTGTGTTATTCAGGAGCTCGGATCTTCGGTGTGAACCTTC 704
 Db 240632 CAGCATTTATTTCCGGTATCTTTTGGTGTAACTGAGCCAGCAATTAACGGGGTTACTTTAC 240573
 QY 705 GCCTGGCTGGCGGTCTTTCATCGGTATCGGTATCGGTACCGCAGTATCGGTGGCGCTTTCATTTG 764
 Db 240572 CACTGAAAAAACAATTCATCATGAGCTGTATCGAGTGTCTGTGGTGGAGGAATCATTTG 240513
 QY 765 CACTCTTTAATATCAAGGAGTGTGGCTTGGGGCTGTCAGGTTTCTTGGGTGTGTTTCTA 824
 Db 240512 GTTTGTAAAGTGCAAAAGTTTATATCATGSGTGAGCTTGGAAATCTTCGCTCTTCTAACT 240453
 QY 825 TTGATGCTCCAGATATGTCATGTTCTTGGTGTGTCAGTGTGTTACTCTTCTTCATCGCAT 884
 Db 240452 TCTTCCAAACAGGTGCTGTGTTTACAA-----GTGCAATCTGTTGGGTAGTTATCGCAA 240399
 QY 885 TCGGGCGCAGCAATGCTTATGGCTTTTACTTGTTCGCGCAGCAACGCGCAGCATTTGATCCAG 944
 Db 240398 TCGTTATCAGCTTCTCTTGTGTTTCACTTAACTTATGTAGCAGGATTCAGAGATCCAG 240339
 QY 945 ATGCAACCGCTGCTCCAGTGCCTGCGAGGAACCAACCAAGCCGAAGCAGAACCCCGCAG 1004
 Db 240338 CTGAAGCAGTTGTAGAAGAAACTAACGTA-----CAGAGGGCG 240300
 QY 1005 AATTTTCAACGATTCACCAATCATCTCAGGCACTTTGACCGGTGAAGTATTTGCACTGA 1064
 Db 240299 AAACTTTAATCGAACGTCAGCAATTCAGCTCCAGTATTCGCGAAATCGTAACTTTAG 240240
 QY 1065 GCAGGCTCAGCGATGCTATGTTTGCAGCGAAAGCTTGGCTCGGCGTTCGATCGCTCC 1124
 Db 240239 CAGAGTAAAGACGAAGCAATTTTCATCGGTGCACTTGGAAAAGGTGAGCAATTTATCC 240180
 QY 1125 CAACCAAGGGCAGTTAGTTTCTCCGGTGAAGTGAAGATTTGTGTGGCATTTCCCATCTG 1184
 Db 240179 CTACTGTGGACGCGTGTGTAGCTCCAGCAGCAGGAACAGTAACTATCTTCCCACTG 240120
 QY 1185 GCCATGCTTTCGAGTTTCGACCAAGCTGAGGATGTTTCCAAATGTGGATATCTTTGATGC 1244
 Db 240119 GTCATGCAATCGGTATTACAAGCAATGA-----CGGTGCAAGATTAATATCC 240072
 QY 1245 ACATTTGGTTTCGACACAGTAAACCTCAACGCGCAGCTTTTAAACCGGTGAGAGACAG 1304
 Db 240071 ACATCGGTATGATACAGTTCACTAGAGTAAGTTCTTTCAGACCATGTTTAAACAG 240012
 QY 1305 CGGATGAAGTCAAAAGCAGGGAGCTGCTGTGTGAATTCGATTTGATGCCATTAAGGCTG 1364
 Db 240011 CGGATGAATCGAAAAAGGTCAATTTACTTAATGAAATTTGATTTGAAGGCATTAAGCAG 239952
 QY 1365 CAGGTTATGAGGTAAACGCGCGAATGTTGTTTTCGAATTAAGAAAAACCGGACCTGTAA 1424
 Db 239951 CTGGATATGATGTTTACAACTCCAGTAGTAGTAAACAACTCCAATCAATATTATAGACGTAA 239892
 QY 1425 ACATTAAGTTTGGCGGAAATTTGAAGCGGAG 1457

QY 1106 TCGGCGGTGGCCATCGTCCCAACCAAGGGCGAGTTAGTTTCTCCGGTGGATGCGAAAGATT 1165
 Db 1519 CAAGGTATCGCTGTGTAACCAAGTCAAGCGGTGCTATGCACTAGCTGATGCTGAAGTT 1578
 QY 1166 GTGGTGGCATTCGCCATCTGCCCATGCTTTTCGCAGTTTCGCCACCAAGGCTGAGGATGTTCC 1225
 Db 1579 TCAATTGCGCTTTCCACAGGGCAGCGCTTTTGGTTTGAAACAAGAAATGG----- 1628
 QY 1226 AATGTGGATATCTTGATGCACATTTGGTTTCGACACAGTAAACCTCAACGCGACGCACTTT 1285
 Db 1629 --TGCTGAAGTTTGTGATTACGTTTGTGATTAATGACACTGTATCTATGAACGCTGACGCTTTT 1686
 QY 1286 AACCGCTGAAGAACAGCGGCGATGAAGTCAAGACAGGGGAGCTGCTGTGTGAATTCGAT 1345
 Db 1687 GAAGCAAAAGTTGCTCAAGGTAATAAGGTGAAAGCTGGCGAAGTTCTTGGAAACATTTGAC 1746
 QY 1346 ATTTGATGCCATTAAAGGCTGCAGGTTATGAGGTAAACCAACGCCGATTTGTTTTCGAAT 1402
 Db 1747 TCAACAATAATCGCTGCAGCTGCACTTGATGATACACAAATGTTATCGTTACAAAT 1803

RESULT 14
 US-10-474-776-152
 ; Sequence 152, Application US/10474776
 ; Publication No. US20040110181A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wyeth
 ; TITLE OF INVENTION: NOVEL STREPTOCOCCUS PNEUMONIAE OPEN READING FRAMES ENCODING POLYH
 ; TITLE OF INVENTION: ANTIGENS AND USES THEREOF
 ; FILE REFERENCE: AM100649-PC7
 ; CURRENT APPLICATION NUMBER: US/10/474, 776
 ; CURRENT FILING DATE: 2003-10-14
 ; NUMBER OF SEQ ID NOS: 752
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 152
 ; LENGTH: 1185
 ; TYPE: DNA
 ; ORGANISM: Streptococcus pneumoniae
 US-10-474-776-152

QY	719	TTCTTCATCGGTATTCGGTACCCAGCTATCCGTGGCGCTTGATGTGACCTCTTTAATAATC	778
Db	454	TTCTTCCTTTCAATTGATTTGCTGGTCAATCCGTGTGGATTGGCTTCTATCCTTGACATT	513
QY	779	AAGGCAGTTGCGTTGGCGCTCGAGTTCTTTGGGTGTGTCTTATTTGATGCTCCAGAT	838
Db	514	GCTGTACTGTGAATGGTATACCATCATCCCTGGTACATGCTTTATGTGTGAACGA	573
QY	839	ATGGTCAATGTTCTTGGTGTGCGAGTTGTACTTCTTCATCCGCAATCGGCGACGAT	898
Db	574	CAACTCCCAATACCTTCTTATGGTAGCTGATCATT-----TGCCCTT	618
QY	899	GCTTATGGCCTTTACTTGGTTCGCGCAACGGCAGCATTCAGATGCACACCGTCTGCT	958
Db	619	GGTTTTGCTCTTACTTTACATGTTTGGTTACG--AAGATGAAGTAGACGCAACTCGACT	675
QY	959	CCAGTGCCTGCAGGACGACCAAGCCGAGCAGAGCACCCGCGAGAAATTTTCAACAGAT	1018
Db	676	GCAAAACGAGCTGAATGGCTGAAGAAAAGAAAGATTGCGCCAGCAGCTCTTCAAAT	735
QY	1019	TCCACCATCATCCAGGCACTTTGACCGGTGAAGCTATTGCACTGAGCAGCGTCAAGCAT	1078
Db	736	GAACAACCTGT---AACTCCTATCGTCGGTGATGTTGTGCTCTTCTGATGTCATGAC	792
QY	1079	GCCATGTTGCCAGCGGAAGCTTGGCTCGGGGTTGCCATCGTCCACCAAGGGCAG	1138
Db	793	CCAGTCTTCTCAAGTGGAGCTATGGGACAAAGTATCGTTGTGAAACCAAGCCAGCGTG	852
QY	1139	TTAGTTTCTCCGCTGAGTCGAAGATTTGTTGGCAATCCCATTCGCCATGCTTTCCGA	1198
Db	853	GTCTATGCACACGCTGATGCTGAAATTCAATTGCCCTTTCCACAGGCGACGCTTTGGT	912
QY	1199	GTTCCGACCAAGGCTGAGGATGTTCCAAATGTGGAATCTTTGANGCACAATGGTTTCGAC	1258
Db	913	TTGAAACACAGAATGTTGCT-----GAAGTTTGTATTGATTTGGTATTTGAT	960
QY	1259	ACAGTAAACCTCAACGCGACGCACTTTAACCCGCTGAAGAGCAGGCGGATGAAGTCAAA	1318
Db	961	ACTGTACTATGAACGGTGACGGTTTGTGAAACAAAAGTTGCTCAAGGTAATAAGGGTAAA	1020
QY	1319	GCAGGGAGCTGCTGTGTGAATTCGATATTGATGCCATTAAAGGCTCAGGTTATAGGTA	1378
Db	1021	GCTGGCGATGTTCTTTGAAACATTTGACTCAAACAAATCGCTGCAGCTGGACTTGAT	1080
QY	1379	ACCACGCGGATTTGTTTTCGAATTACAAGAAAACCGGACCTGTAAACACTTACCGTTTG	1438
Db	1081	ACAACAATGTTATCGTTACAATACAGGTGACTACGCTTCAGTAGCTCCAGTCGCAACA	1140
QY	1439	GGCGAAATGAAGCGGAGCGCAACCTGCTCAACGTGCGCAAGAAA	1483
Db	1141	GGTTCAGTTGCTTAAGGGGGATGCTGTGATCGAAGTGAAATCTTAA	1185

RESULT 15

```

US-08-961-527-30/c
Sequence 30, Application US/08961527
Publication No. US20020032323A1
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
NUMBER OF SEQUENCES: 391
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text

```

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/961,527
;; FILING DATE:
;; CLASSIFICATION: 424
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER:
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Brookes, A. Anders
;; REGISTRATION NUMBER: 36,373
;; REFERENCE/DOCKET NUMBER: PB340P1
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (301) 309-8504
;; TELEFAX: (301) 309-8512
;; INFORMATION FOR SEQ ID NO: 30:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 9769 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
US-08-961-527-30

Query Match 7.1%; Score 107.8; DB 8; Length 9769;
Best Local Similarity 46.6%; Pred. No. 4.1e-22;
Matches 552; Conservative 0; Mismatches 597; Indels 36; Gaps 5;

QY 302 GACTTCCTGATGCTCCAGTGTGAGCTTCTGCTCAACCGGATTCCTTCAATTCATCGCC 361
DB |||||
QY 362 ATGGGCCAGCATCGCTGGTGGGCGATGCTGGCACACGGCTTACAGGACATTTAT 421
DB |||||
QY 8919 ATTGGACCGATTTCACGTTGTTGAAACTACATCCTTATTGCTACAAAGCGATTCT 8860
QY 422 GATTTCGGTGGTCCAGTCCGGCGCTGCTCTTCGGTCTGGTCTACTACCAATCGTCA 481
DB |||||
QY 8859 AGCATGCCATTGGTCTTGGTGGTCTTCTGATTTGGTGGGTTCCACCAATGATGCTGG 8800
QY 482 ACTGTCGTGACAGTCTTCCGCCCAATGAGCTGGAGCTGTTTAAACAGGG---TGGA 538
DB |||||
QY 8799 TCAGGTGTGCACCATCTTCAACTTGTGTAAGTGCATTTACTTCTGCTGACCAATGCT 8740
QY 539 TCCTTCATCTTCGACGCACTATGCTGATATATCGCCAGGTCGGCATGTTTGGCA 598
DB |||||
QY 8739 AACCCATTCAACGCTATCATCAAGCTGCTATGACAGCTCAAGTGTCTGCTACTGTTGG 8680
QY 599 GTGTTCTTCCTGGCGAAGAGTGAAAGCTCAAGGGCTTGCAGGTGCTTTCAGGTGCTCC 558
DB |||||
QY 8679 GTTGGTGTAAACAAAATCCAAACTGAAACACCTTGTCTCCCGCTGCTCTTTCT 8620
QY 659 GCTGTTCTTGGTATTACGGAGCTGCGATCTTCGGTGTGAACCTTCGCTCGCTGGCGG 718
DB |||||
QY 8619 GCCTTCCTAGGTATTACAGAGCTGCTATCTTCGGGGTGAACCTTGGCTTCGGTAAACCA 8560
QY 719 TTCTTCATCGGTATCGGTACCGCAGTATCGTGGCGCTTGAATGCACTCTTTAATATC 778
DB |||||
QY 8559 TTCTTCCTTTCAATGATTGCTGGTGCATTCGCTGGTGGATTTGGCTTCTATCCTTGA 8500
QY 779 AAGGCAGTTGCTTGGCGCTGCTGAGTTTCTTGGGTGTTGTTCTATTGATGCTCCAGAT 838
DB |||||
QY 8499 GCTGTAAGTGAATGATATACCATCATCCCTGGTACAACTCTTATGTTGGTAAACGA 8440
QY 839 ATGTCATGTTCTTGGTGTGTCAGTTGTTTACTTCTTCATCGCATTCGGCGCAGCGATT 898
DB |||||
QY 8439 CAACATCCACATATCTTATTGTTAGTGTATCATTT-----TGCCCTT 8395
QY 899 GCTTATGCTTTTACTTGGTTCGGCGCAACGCGCAGCATTTGATCCAGATGCAACCGCTCT 958
DB |||||
QY 8394 GGTTTGCTCTTACTTACATGTTTGGTTACG---AAGATGAAGTAGACGCAACTCAGCT 8338
QY 959 CCAGTGCCTGAGGAACGACCGAAGCCGAGCAGACCGCAGCAATTTTCAACCAT 1018
DB |||||
QY 8337 GCATAACGAGCTGAAGTGGTGAAGAAAAGAAAGTGGCCAGCAGCTCTTCAAAAT 8278

Search completed: October 30, 2004, 18:51:03
Job time : 760 secs

QY 1019 TCACCATCATCCAGGACACCTTTGACCGGTGAAGCTATTGCACTGAGCAGCGTCAAGCAT 1078
DB |||||
QY 8277 GAAACACTTGT---AATCTTATCGTGGTGAATGTTGTGCTCTTGTGATGCAATGAC 8221
DB |||||
QY 1079 GCCATGTTTCCAGCGGAAAGCTTGGCTCGGGCGTTGCCATCGTCCCAACCAAGGGGCGAG 1138
DB |||||
QY 8220 CCAGTCTTCTCAAGTGGAGCTATGGGACAAGGTATCGTTGTGAACCAAGCAAGCGCGTG 8161
QY 1139 TTAGTTTCTCGGTGAGTGGAAAGATTGTGGTGGCATTCCCATCTGGCCATGCTTTTCGCA 1198
DB |||||
QY 8160 GTCTATGCACCGAGCTGATGCTGAAGTTTCAATTGSCCTTTCCAAACAGGGCAGCTTTTGGT 8101
QY 1199 GTTCGCACCAAGCTGAGGATGGTTCCAAATGTGGATATCTTGATGCACTTGGTTCGAC 1258
DB |||||
QY 8100 TTGAAAACAAGAAATGGTGCT-----GAAGTTTGTGATTCATGTTGGTATTGAT 8053
QY 1259 ACAGTAAACCTCAACGGCACCGCACTTTTAAACCGCTGAAGAGCAGGGCGATGAGTCAAA 1318
DB |||||
QY 8052 ACTGTATCTATGAACCGGTGACGGTTTGAACAAAAGTTGCTCAAGGTAATAAGGTGAAA 7993
QY 1319 GCAGGGGAGCTGCTGTGTGAATTCGATATTGATGCCATTAAAGGCTGCAGGTATGAGGTA 1378
DB |||||
QY 7992 GCTGGCGATGTTCTTGGAAACATTTGACTCAACAAAATCGCTGCAGCTGCACTTGATGAT 7933
QY 1379 ACCACGCCGATTTGTTTGGAAATTACAAGAAACCGGACCTGTAAACACTTACGGTTTG 1438
DB |||||
QY 7932 ACAAATGTTTATCGTTTACAAATACAGGTGACTACGCTTTCAGTAGCTCCAGTCGCAACA 7873
QY 1439 GCGGAAATTGAACGGGAGCCAACTGCTCAACGTCGCAAGAAA 1483
DB |||||
QY 7872 GGTTCAGTTGCTAAGGGGATGCTGTGATCGAAGTGAATAATCTAA 7828

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 30, 2004, 13:44:46 ; Search time 4437 Seconds
(without alignments)
12540.787 Million cell updates/sec

Title: US-09-604-231-1
Perfect score: 1527
Sequence: 1 ctcatggcctctgcgcgtt.....gttgaaccttgagtgttgcg 1527

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gssi:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query % Match	Length	DB ID	Description
C 1	105.4	6.9	917	8	BZ685850
C 2	92.6	6.1	750	9	CL686986
C 3	80.8	5.3	770	7	CN823265
C 4	73.2	4.8	861	9	CL657583
C 5	63.2	4.1	1412	9	CG744886
C 6	62	4.1	788	9	CL693618
C 7	48.8	3.2	680	1	AL506262
C 8	49.8	3.2	775	9	CL675911
C 9	47.4	3.1	1328	8	BZ572810
C 10	46.4	3.0	763	9	CL695497
C 11	45.8	3.0	473	7	TO3017
C 12	45.4	3.0	694	4	EM100823
C 13	45.4	3.0	791	9	CL661694
C 14	44.4	2.9	785	6	CD443326
C 15	43.2	2.8	869	7	CK159167
C 16	42.6	2.8	473	9	CE077138
C 17	42.2	2.8	798	6	CD780937
C 18	42	2.8	750	4	BM559084
C 19	41.8	2.7	681	9	CNS02EOD
C 20	41.8	2.7	889	7	CK159613
C 21	41.6	2.7	895	9	CNS0071A
C 22	41.4	2.7	477	2	BE997958
C 23	41.4	2.7	518	6	CA920257
C 24	41.4	2.7	547	6	CD442889
C 25	41.4	2.7	601	4	BG580528
C 26	41.4	2.7	627	2	BE998413
C 27	41.4	2.7	793	6	CF069882
C 28	41.4	2.7	511	5	BU345319
C 29	41.2	2.7	581	5	BQ988486
C 30	41	2.7	683	5	BQ986937
C 31	41	2.7	732	5	BQ933574
C 32	41	2.7	739	5	BQ991717
C 33	41	2.7	744	5	BQ988795
C 34	41	2.7	745	5	BQ986791
C 35	41	2.7	759	5	BQ993437
C 36	41	2.7	776	5	BQ995038
C 37	41	2.7	884	9	CNS006UO
C 38	40.8	2.7	983	9	CNS03CVD
C 39	40.8	2.7	1101	9	CNS00LXJ
C 40	40.8	2.7	970	9	CNS03H6V
C 41	40.2	2.6	902	8	CC366418
C 42	40	2.6	548	7	CK895185
C 43	39.8	2.6	606	2	BE997957
C 44	39.8	2.6	714	7	CF753589
C 45	39.6	2.6			

ALIGNMENTS

RESULT 1
BZ685850/c
LOCUS PUBRE66TD ZM 0.6 1.0 KB Zea mays genomic clone ZMMBTa035K11,
DEFINITION genomic survey sequence.
ACCESSION BZ685850
VERSION BZ685850.1 GI:28245462
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 917)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
Maize Genomics Consortium
Unpublished (2003)
CONTACT: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@cigr.org
Seq primer: 1P
Class: sheared ends.
Location/Qualifiers
1. 917
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBTa035K11"
/clone_lib="ZM 0.6 1.0 KB"
/note="Vector: pCR4-TOPO; Site:1: EcoRI; 0.6-1.0 kb high
CoT selected genomic DNA library"

FEATURES
source

Query Match 6.9%; Score 105.4; DB 8; Length 917;
Best Local Similarity 52.6%; Pred. No. 1.2e-18;
Matches 308; Conservative 0; Mismatches 261; Indels 16; Gaps 3;
QY 1 CTATGCGCATCTGCGCGTTCGGCTTTTCCAGTGTGGTGTTCACCGCAACCAAG 60
|||||
DB 585 CTGATGCCAACACCTCGTTCGTCTCTCCACGCGTGGTGGCTGGTGGCGGCGCAAG 526
|||||

/clone_lib="Sheep spleen\brain pSport1 library"

```

ORIGIN
Query Match      5.3%; Score 80.8; DB 7; Length 770;
Best Local Similarity 50.4%; Pred. No. 1.4e-11;
Matches 259; Conservative 0; Mismatches 242; Indels 13; Gaps 2;

QY 1 CTATGCACTCGCGCTTTCGCGTTCCTCCAGTGTGGTTGCTTTCACCGCAACCAAG 60
Db 258 CTATGCCCAACACCTCGTTCGTTCCTCCGCGCTGTGGTGGTGGCGGCAAG 317

QY 61 CGTTTCGGCGCAATGAGTTCTCGGCGCGCGCTATTTGTTATGGCATGTGTTCCCGAG 120
Db 318 CGTTTGGTGGCATGAAATCTCGGCATGCTCC-TGGGCTCATGCTGTGTACACCCCGA 376

QY 121 CTTGGTGAACGGCTACGACCTGGCGCGCACCATCTGGTCTCGGGCGAAATGCCATGTGTC 180
Db 377 TTATGCTCAACGCTTGAACCTACGCAAGGCGTCCGCGGCTCGACGGCCAGAGCTGCC 436

QY 181 CCTGTTTGGT-----TTAGATGTGTCCTCCAGCGGTTCACAGGGCACCGTGCT 228
Db 437 GTATCTTCGATATTTTCGGCTGGTTCAGATCGAAGAGTCCGGTACCCAGGGCGAGATCCT 496

QY 229 TCCTGTGCTGTGGTTCCTTGGATTCTGSCAACGATCGAGAGTTCCTGCTCACCGGATTCCT 288
Db 497 GCGATCTCTGCTGGCGCTACGTGAAGCGTGAATGAAATGCTCGGGCGCGGT 556

QY 289 CAAGGCACTGACAGCTTCTGATCACTCAGTGTGAGTGTCTGCTCACCGGATTCCT 348
Db 557 GCGCATGCGCATCAATCTGCTGTGTCGGATCACCACCATCTGCTCACTGGCGTGT 616

QY 349 TACATTCATCGCATGCGCCAGCAATGCTGGTGGCGCATGTCTGGCAACAGTCT 408
Db 617 GCGCTGGCGATCATCGGCGCGGTGACCGGTACCTCGGCATCTGATCACCGAGGCT 676

QY 409 ACAGGACATTTATGATTTTCGGTGTCCACTCGCGGCTGCTCTGCTGTCTGTCTACTC 468
Db 677 GGTACTCTGTTGACCTGGCGCCCATGCTCGCGGTGCTCAATTCGGCGCTGTGTACGC 736

QY 469 ACCAATCGTCACTGCTGTCTGACCGAGTCCTTC 502
Db 737 ACCGCTGGTGATCACCGGATGACCACTATTC 770

RESULT 4
CL657583
LOCUS
DEFINITION
CL657583.1 GI:50139334 861 bp DNA linear GSS 09-JUL-2004
PR10129d.H09 - PR10129d.B21 (861) Mixed stage fosmid library of P.
pacificus var. California Pristionchus pacificus genomic, genomic
survey sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Pristionchus pacificus
Pristionchus pacificus
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
Neodiplogasteridae; Pristionchus.
1 (bases 1 to 861)
REFERENCE
Srinivasan,J., Octo,G.W., Kahlow,U., Geisler,R. and Sommer,R.J.
AppaDB: an AcedB database for the nematode satellite organism
Pristionchus pacificus
Nucleic Acids Res. 32 (1), D421-D422 (2004)
Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel.: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@tuebingen.mpg.de
This library was generated at Caltech, Pasadena, USA and end
sequenced at Vancouver, Canada.
Seq primer: T7
Class: fosmid ends.

FEATURES
source
Location/Qualifiers
1..861
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="Mixed stage fosmid library of P. pacificus
var. California"
/notes="Vector: pEpifos-5 Fosmid vector"

ORIGIN
Query Match      4.8%; Score 73.2; DB 9; Length 861;
Best Local Similarity 55.0%; Pred. No. 2.3e-09;
Matches 144; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

QY 1213 TGAGATGGTTCATATGTGATATCTTGATGCACTTGTGTCACACAGTAACCTCAA 1272
Db 86 TGAGTCAGATGATGCTGTGAGATCCTGATTCATGTCGTTATCGACACCGTAACCTGGA 145

QY 1273 CGGCACGCACTTAAACCGCTGAAGAGCAGGCGATGAAGTCAAGCGGGAGCTGCT 1332
Db 146 CGGCAAAATCTTTTCGCTCAACGTCAACGTGGGTGACAAGGTCAATACAGCGCATCGCT 205

QY 1333 GTGTGAATTCGATATGATCCATTAAAGCTGCAGGTATAGGTAAACCAACCGCGATTCT 1392
Db 206 GATTTCTTTGATATCCCTGCTATTTCGCGAGCGCGATTGATGACACCGCGGTATT 265

QY 1393 TGTTTGAATTCAGAAACCGACCTGTAAACACTTACGTTTGGGCGAAATTTGAAGC 1452
Db 266 AATCAGTATAGCATGATTTTACAGACGTATTACCCACGCGACGCGCAGATAGCGC 325

QY 1453 GGGAGCCACCTGCTCAACGTC 1474
Db 326 AGTGAAACCGCTGTTATCCATC 347

RESULT 5
CG744886/c
LOCUS
DEFINITION
CG744886
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Pristionchus pacificus
Pristionchus pacificus
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
Neodiplogasteridae; Pristionchus.
1 (bases 1 to 1412)
REFERENCE
Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,
Buntjer,J., van der Meulen,M. and Sommer,R.J.
An integrated physical and genetic map of the nematode Pristionchus
pacificus
Mol. Genet. Genomics 269 (5), 715-722 (2003)
12884007
Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel.: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@tuebingen.mpg.de
Class: BAC ends.

FEATURES
source
Location/Qualifiers
1..1412
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="Ppa EcoRI BAC Library"
/notes="The library was generated by a partial digest of
the genomic DNA with EcoRI and cloning into the BAC

```


vector."

ORIGIN

Query Match 4.1%; Score 63.2; DB 9; Length 1412;
Best Local Similarity 46.6%; Pred. No. 2.1e-06;
Matches 274; Conservative 0; Mismatches 308; Indels 6; Gaps 2;
172 AATGTGTCCTGCTTTGGTTAGATGTTGCGCAAGCGGTTACAGGCGACCGTCTTC 231
Db |||||
655 AGTGTGGCATTTGGCATGTTTCAGCATCGGCAAGTGGCTATCAGGCACAGGTGATCCC 596
Qy |||||
232 TGTGCTGGTGTTCCTTGGATTCTGGCAAGCATCGAGAGTCTCTGCAACAGGACTCAA 291
Db |||||
595 GGCACCTTTAGCCGGGCTGGCAGCTGGGCGTTATTGAACTCGCCTTAAACGATCGTACC 536
Qy |||||
292 GGGCACTGCAGACTTCCTGATCACTCCAGTGTCTGACGTTGTGTCTACCGGATTCCTTAC 351
Db |||||
535 GGATTACCTCTATCTGGTGGTGGTGGCGTCTGTGCTGATCTCGCGGTGTCTCTCGC 476
Qy |||||
352 ATTCAATCGCCATTCGGCCAGCAATGCGCTGGTGGGCGATGTGTGGCACACGCTCTA-- 409
Db |||||
475 CCATGCGCTGATTGGTTCGGTTTGGTCGCATGATTGGCGATGGCGTTCGCTTTGGCGTACG 416
Qy |||||
410 -CAGGACTTTATGATTTTCGCTGTGCTCAGTCCGCGTCTGCTCTTGGTCTGTCTACTC 468
Db |||||
415 TCACCTGATGACCGGCGAGCTTGTCTCCGATTGGTGGCGCATGTTTGGCTCTCTGTAGGC 356
Qy |||||
469 ACCAATGCTCATCTGCTGTCACAGTCTCTCCGCAATTCAGCTTGGAGCTGTGTTT-- 526
Db |||||
355 GCGCTGTGATCACCGCGGTACACAGCACACCCCTTCTATTGATTGGCAGATGATTCA 296
Qy |||||
527 -AACGAGGTGGATCCTTCACTTCGCAAGCGCATCTATGGCTATATCGCCAGGTGC 585
Db |||||
295 AAGCATGGTGGACCGCAGTGGCGCTGATTGCGCTGTGCAATATCGCTCAGGGCTC 236
Qy |||||
586 GGCATGTTTGGCAGTGTCTTCTCGCGAAGTCAAAAGCTCAAGGCGCTTCGAGGTGC 645
Db |||||
235 GCGCTGATGAGCATATCAATTCAGCGCGGAGCAATGACATGACGAGATCTCCGTGCC 176
Qy |||||
646 TTCAGTGTCTCGGCTGTTCTTTGGTATTACGAGCGCTCGATCTTCGGTGTGAACCTCG 705
Db |||||
175 TGCCGCTATCTCCGCTGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 116
Qy |||||
706 CTGCGCTGCGCTTCTTCATCGTATCGGTACCGCAGTATCGGTGG 753
Db |||||
115 ATATGCTCTCCGATGCTGTGGCGCATGATTGTTCTGCTGTGGCAGG 68

RESULT 6
CL693618 788 bp DNA linear GSS 10-JUL-2004
LOCUS PRI0162a_D06_2 - PRI0162a.BR (788) Mixed stage fosmid library of P.
pacificus var. California Pristionchus pacificus genomic, genomic
survey sequence.
CL693618
CL693618.1 GI:50215526
GSS.
Pristionchus pacificus
Pristionchus pacificus
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
Neodiplogasteridae; Pristionchus.
1 (bases 1 to 788)
Srinivasan,J., Otto,G.W., Kahlou,U., Geisler,R. and Sommer,R.J.
AppaDB: an AcedB database for the nematode satellite organism
Pristionchus pacificus
Nucleic Acids Res. 32 (1), D421-D422 (2004)
Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@tuebingen.mpg.de

FEATURES
source

Location/Qualifiers
1..788
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="Mixed stage fosmid library of P. pacificus
var. California"
/note="Vector: pPifos-5 Fosmid vector"

ORIGIN

Query Match 4.1%; Score 62; DB 9; Length 788;
Best Local Similarity 52.3%; Pred. No. 3.8e-06;
Matches 137; Conservative 0; Mismatches 125; Indels 0; Gaps 0;
1229 GTGATATCTTGTATGCACTTGGTTTCGACACATTAACCTCAACGGCAGCATTAAAC 1288
Db |||||
388 GCGGAGATCGTCGTCATATGGGTATCGACACCGCTAGCGCTGGAAGGTAAAGGCTTTAAA 447
Qy |||||
1289 CCGCTGAAGAGCGAGGCGATGAAGTCAAGCAGGAGGCTGCTGTGTAATTCGATATT 1348
Db |||||
448 CGTCTGTTGAAGAGGGGGCGCAGGTAAGCGGCAACCGATTCTGGAATGATCTG 507
Qy |||||
1349 GATGCCATTAAAGCTGCGAGGTTATGAGTACACCGCGATTGTTTTCGAATTACAAAG 1408
Db |||||
508 GATTACCTGGAACGCTTAACGCCCGCTCGATGATTAGCCGGTGGTTTCGAGCAATATCGAC 567
Qy |||||
1409 AAAACCGGACCTGTAACACTTACCGTTTGGCGCAATTAAGCGGAGCAGCACTGCTC 1468
Db |||||
568 GATTTCAGTGGCTTGATCATTAAGCTCAGGCGCATGTTGTGGCGGTCAACACCGCTG 627
Qy |||||
1469 AACGTCGCAAGAAAGAGCGG 1490
Db |||||
628 TATGAATCAAAAAGTAATCTG 649

RESULT 7
AL506262/c

AL506262 680 bp mRNA linear EST 04-JAN-2001
LOCUS Hordeum vulgare Barke developing caryopsis (3.-15.DAP)
DEFINITION Hordeum vulgare subsp. vulgare cDNA clone HY02124T 5', mRNA
sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AL506262
Hordeum vulgare subsp. vulgare
sequence.
AL506262.1 GI:12032477
EST.
Hordeum vulgare subsp. vulgare
Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poaceae; Tricaceae; Hordeum.
1 (bases 1 to 680)
Michalek,W., Weschke,W., Pleissner,K.-P. and Graner,A.
EST sequencing and analysis in barley
Unpublished (2000)
Contact: Michalek W
Institute for Plant Genetics and Crop Plant Research
Corrensstr.3, D-06466 Gatersleben, Germany
Email: michalek@ipk-gatersleben.de, http://pgrc.ipk-gatersleben.de
Seq primer: T3 primer for 5'end.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 680)
Location/Qualifiers
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="barke"
/sub_species="vulgare"
/db_xref="taxon:112509"
/clone="HY02124T"
/tissue_type="developing caryopsis (3.-15.DAP)"
/lab_host="XL0LR"

FEATURES
source

Location/Qualifiers
1..680
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="barke"
/sub_species="vulgare"
/db_xref="taxon:112509"
/clone="HY02124T"
/tissue_type="developing caryopsis (3.-15.DAP)"
/lab_host="XL0LR"

BZ572810
LOCUS
DEFINITION msh2_2811.x3 msh Pseudomonas aeruginosa genomic clone msh2_2811,
genomic survey sequence.
ACCESSION BZ572810
VERSION BZ572810.1 GI:27207871
KEYWORDS GSS.
SOURCE Pseudomonas aeruginosa
ORGANISM Pseudomonas aeruginosa
REFERENCE
AUTHORS Spencer,P.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
Burns,J.B., Kaul,R. and Olsen,M.V.
TITLE Whole-Genome-Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
JOURNAL J. Bacteriol. (2002) In press
COMMENT Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 20622216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.

FEATURES
source
Location/Qualifiers
1..1328
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="MSH"
/db_xref="taxon:287"
/clone_lib="msh2_2811"
/clone="msh2_2811"
/note="Environmental isolate. Whole genomic shotgun
library."

ORIGIN
Query Match 3.1%; Score 47.4; DB 8; Length 1328;
Best Local Similarity 49.2%; Pred. No. 0.074; 156; Indels 12; Gaps 1;
Matches 163; Conservative 0; Mismatches 156; Indels 12; Gaps 1;
QY 1027 CATCAGGCACCTTTGACCGGTGAAGCTATTGCACTGAGCAGCGTCAGCGATGCCATGTT 1086
DB 253 CTTCAAGCGCCCTGAGCGGCCCGGTGATGCGGCTCAACGAGTACCGACCGGTGTT 312
QY 1087 TCCAGCGGAAGCTTGGCTCGGCGTTCGATCGTCCCAACCAAGGGCAGTTAGTTTC 1146
DB 313 TTCCAGCGGCACCTTGGCGGAAGGATCGCCATCGACCCGCTCAACGACTGCTGCACGC 372
QY 1147 TCCGGTGTGGAAGATGTGTGGCATTCCTCCATCGGCCATGTTTCGACGTTTCGCAC 1206
DB 373 GCCCTGCGCGCGGTGGTCAGCCACTTGGCGCGAACCCTCCAGCGCTCAGCTTGGCGC 432
QY 1207 CAAGGCTGAGGATGTTCCAAATGTGGATATCTTGATGCACATTTGGTTTCACACAGTAA 1266
DB 433 CGACACGG-----TGCCGAGCTGTGCTGCAGCTCGGCTCGACACCGTACA 480
QY 1267 CTTCAACGGCAGCATTTAAACCGCTGAAGAGCAGGCGGATGAGTCAAGACGGGA 1326
DB 481 GCTGAAGGCGAAGGCTTCGAGCGCTTGTTCGAGGAAGGTGCGCGGTGATCGAGGGCCA 540
QY 1327 GCTGCTGTGAATTCGATATTGATGCCATT 1357
DB 541 GCGCTGCTGCGCTTCGACTGACCGCGTT 571

RESULT 10
LOCUS CL695497 763 bp DNA linear GSS 10-JUL-2004
DEFINITION PR1016c_C08_2 - PR1016c.BR (763) Mixed stage fosmid library of P.
pacificus var. California P. Pristionchus pacificus genomic, genomic
survey sequence.
ACCESSION CL695497

VERSION
KEYWORDS
SOURCE
ORGANISM

CL695497.1 GI:50217405

GSS.
Pristionchus pacificus
Pristionchus pacificus
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
Neodiplogasteridae; Pristionchus.

REFERENCE
AUTHORS Srinivasan,J., Otto,G.W., Kahlow,U., Geisler,R. and Sommer,R.J.
TITLE AppADB: an AcedB database for the nematode satellite organism
JOURNAL Nucleic Acids Res. 32 (1), D421-D422 (2004)
COMMENT Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@tuebingen.mpg.de
This library was generated at Caltech, Pasadena, USA and end
sequenced at Vancouver, Canada.
Seq primer: T7
Class: fosmid ends.

FEATURES
source
Location/Qualifiers

1..763
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="Mixed stage fosmid library of P. pacificus
var. California"
/note="Vector: pSpifos-5 Fosmid vector"

ORIGIN

Query Match 3.0%; Score 46.4; DB 9; Length 763;
Best Local Similarity 45.6%; Pred. No. 0.12;
Matches 247; Conservative 0; Mismatches 286; Indels 9; Gaps 2;
QY 224 GTGCTTCCTGCTGGTGGTTCCTGGATTCCTGGATCTCTCCAGTGTGACGTTCTCTGACAG 283
DB 1 GTGTTCCTCCGCGCTGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 60
QY 284 CGACTCAAGGCACCTGACAGCTTCCTGATCACTCCAGTGTGACGTTCTGCTCACCAG 343
DB 61 ATTACACCGCGGTGACCAAAACTTCCTCAAGCCAAATGCTGATGTGTGATTCGCGCA 120
QY 344 TTCTTACATTCATCGCCATTGGCCAGCAATGCGCTGGGTGGGCGATGTGCTGGCACAC 403
DB 121 CCGCTGGCAATCCTGCTGATTTGGCCCGATTTGGTATCTGGATCGGTAGCGCCATTTCCGG 180
QY 404 GGTCTTACAGGCACTTTATGATTTGGTGGTTCAGTCCGCGGTCTGCTCTTCTGCTCTGCTC 463
DB 181 CTGGTTTACACCATTCATGGTTATCTGGCTGGCTTTTCACTGCCATTTATGGCGCGCTG 240
QY 464 TACTCACCAATCGTCACTGCTGTCACCAAGTCTTCCCGCCAAATTTAGCTGGAGCTG 523
DB 241 TGCGCTCTGCTGTAATGACCGGATGACCGCGTCTTTTACGCCAACCAATCAATTCAGACC 300
QY 524 TTTTACCGAGGGTGGATCTTTCATCTTCGCAACGGCATCTATG-----GCTAATATCGCC 577
DB 301 ATTTCGGAACCGGCAAGAGAGGATGGTTCATGCGCTCATAGATCGGCGCTAACCTGTG 360
QY 578 CAGGGTGGGCAATGTTTGGCAGTGTTCCTTCCTGGGGAAGAGTGAAGTCAAGGCGCTT 637
DB 361 CTGGCGGTTCATCACTGGCGGTGGGGAAGAAACCGGAACTTCGCCACAGC 420
QY 638 GCAGGTGCTTCAGGTGCTCCGCTGTTCTT---GGTATTACGAGGCTTCGATCTTCGGT 694
DB 421 GCGTGGCTGCGCGGCAATGACCAATATGCGGGGATTTCCGAACCGCGGTATACGGC 480
QY 695 GTGAACCTTCGCTCGCTGGCGGTTCCTTCATCGGATTCGTTACCCAGCTATTCGGTGGC 754
DB 481 GTGGCATCCGCTGAAACGTCGCTTATCGCGCTTATCGCGAGTCTCATCAGCGGCTTTATTGCGGC 540

QY 755 GC 756
Db 541 GC 542

RESULT 11
T03017/c
LOCUS
DEFINITION FB20C8 Fetal brain, Stragatene Homo sapiens cDNA clone FB20C8, mRNA
ACCESSION T03017
VERSION T03017.1 GI:314258
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 473)
Khan,A.S., Wilcox,A.S., Polymeropoulos,M.H., Hopkins,J.A., Stevens,T.J., Robinson,M., Orpana,A.K. and Sikela,J.M.
TITLE Single pass sequencing and physical and genetic mapping of human CDNA5
JOURNAL Nat. Genet. 2, 180-185 (1992)
MEDLINE 94258200
PUBMED 1345165
COMMENT On Sep 21, 1992 this sequence version replaced gi:279156.
Contact: Sikela JM
Department of Pharmacology
University of Colorado Health Sciences Center
Box C236, 4200 E. 9th Ave, Denver CO 80262-0236
Tel: 3032708637
Fax: 3032707097
Email: nikki@tally.uchsc.edu.

FEATURES
source
1. .473
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="ATCC (inhost):86662"
/db_xref="GDB:D0S2785E"
/db_xref="taxon:9606"
/clone="FB20C8"
/clone_lib="Fetal brain, Stragatene"
/note="Vector: Lambda ZAP II; The FB library (catalog #937201, Stragatene) was constructed by directional cloning and oligo d(T)-priming in Lambda ZAP II phage utilizing mRNA from a 17 week human fetus (total brain)."

ORIGIN
Query Match 3.0%; Score 45.8; DB 7; Length 473;
Best Local Similarity 59.7%; Pred. No. 0.15;
Matches 77; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 1231 GGATATCTTGATGCAATGTTTCGACACAGTAAACCTCAACGCCACGCACCTTTAAACC 1290
Db 275 GGAGATCGTCGTCATAGGGTATCGACACCGTAGCGTGGAGTAAAGCGCTTTAAACG 216

QY 1291 CTTGAAGACGAGCGGATGAAGTCAAGACGAGGAGCTGCTGTGTAATTCGATATGA 1350
Db 215 TCTGGTGAAGAGGGTGGCGAGGTAAGCGCAGGCAACCGATCTCGGAATCGATCTGA 156

QY 1351 TGCCATTAA 1359
Db 155 TTACCTGAA 147

RESULT 12
BM100823/c
LOCUS
DEFINITION EBpi01_SQ002_B17 R pistil, 1 DPA, no treatment, cv Optic, EBpi01
Hordeum vulgare subsp. vulgare cDNA clone EBpi01_SQ002_B17 5', mRNA
ACCESSION BM100823
VERSION BM100823.2 GI:21943888

QY 755 GC 756
Db 541 GC 542

RESULT 13
CL661694/c
LOCUS
DEFINITION CL661694
PRI013c.B10 - PRI013c.B21 (791) Mixed stage fosmid library of P.
pacificus var. California Pristionchus pacificus genomic, genomic
survey sequence.

KEYWORDS EST.
SOURCE Hordeum vulgare subsp. vulgare
ORGANISM Hordeum vulgare subsp. vulgare
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Hordeum. 1 (bases 1 to 694)
AUTHORS Hedley,P., Liu,H., Caldwell,D., McCallum,N., Mudie,S., Cardle,L., Ramsay,L., Machray,G., Marshall,D.F.M. and Waugh,R.
TITLE Development of Barley Transcriptome Resources
JOURNAL Unpublished (2001)
COMMENT On Nov 21, 2001 this sequence version replaced gi:17031891.
Contact: Waugh R, Marshall DF
Genome Dynamics/Computational Biology
Scottish Crop Research Institute
Invergowrie, Dundee, DD2 5DA, Scotland, UK
Tel: 00 44 1382 562731
Fax: 00 44 1382 562426
Email: est@scri.sari.ac.uk
All sequence has a Phred quality score of 20 or over
Seq primer: M13 reverse.

FEATURES
Location/Qualifiers
source
1. .694
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="Optic"
/sub_species="vulgare"
/db_xref="taxon:112509"
/clone="EBpi01_SQ002_B17"
/tissue_type="pistil"
/dev_stage="1 DPA"
/lab_host="DH10B"
/clone_lib="pistil, 1 DPA, no treatment, cv Optic, EBpi01"
/note="Vector: pSPORI; Site 1: Sal I; Site 2: Not I; Non-normalised library, directionally cloned into pSPORI. Derived from pistils dissected from developing grains (24 hours post anthesis) in glasshouse grown barley plants. Developed as part of the barley transcriptome resources of BBSRC/SEERAD funded cereal IGF (Investigating Gene Function) project."

ORIGIN
Query Match 3.0%; Score 45.4; DB 4; Length 694;
Best Local Similarity 48.6%; Pred. No. 0.22;
Matches 124; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

QY 636 TTGCAGGTGCTTCAGGTGCTCCGCTGTTCTGGTATTACGAGCCTCGATCTCGGTG 695
Db 693 TTGTTGTGTTCTTCGTCGTGTCGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 634

QY 696 TCACCTTCGCTCGCGTGGCGGTTCTTCATCGGTATCGGTACCGACATPCGTTGCG 755
Db 633 TCATGCTATTGTTTTCGCGAGTTTCGTCGTTGTTGTTGTTGTTGTTGTTGTTGTCG 574

QY 756 CTTTGATTGCACTCTTTAATATCAAGCAGTTCGTTGGCGCTGCAGTTCCTGGGTG 815
Db 573 TTGCTATTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 514

QY 816 TTGTTTCATTGATGCTCCAGATATGGTCATGTTCTTGGTGTGTCAGTTGTTACCTTCT 875
Db 513 TTGTTATAGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 875

QY 876 TCATCGCATTCGCG 890
Db 453 TCATTGTTATCGTCG 439

RESULT 13
CL661694/c
LOCUS
DEFINITION CL661694
PRI013c.B10 - PRI013c.B21 (791) Mixed stage fosmid library of P.
pacificus var. California Pristionchus pacificus genomic, genomic
survey sequence.

Bioinformatics

University of Saskatchewan, Department of Computer Science
1C101 Engineering Building, 57 Campus Drive, Saskatoon,
Saskatchewan, S7N 5A9, Canada

Tel: 306 966 1769

Fax: 306 966 2033

Email: fgas_est@cs.usask.ca

This sequence is the direct result of the Base calling software
Phred (default parameters). It is the raw base calls. To aid in the
identification of the high quality insert the software Lucy
(default parameters) has been run on this sequence. Lucy identified
the region [128,636].

Plate: Talet537 row: N column: 23.

FEATURES

source

Location/Qualifiers

1..869

/organism="Triticum aestivum"

/mol_type="mRNA"

/cultivar="Wheat line PI 178383"

/db_xref="taxon:4565"

/lab_host="DH5 alpha"

/clone_lib="Triticum aestivum FGAS: Talet5"

/note="Organ: Crown; Vector: pGEM-T; SSH (suppression

subtractive hybridization) cDNA library from genotype

PI178383 cold hardened at 2 C for 21 days and 49 days

(equal amount of cDNA pooled together before subtraction,

tester) and subtracted against genotype Norstar cold

hardened at 2 C for 1 day (24 H) (driver). Modified Smart

cDNA (Clontech) priming and non-directional cloning"

ORIGIN

```

Query Match      2.8%; Score 43.2; DB 7; Length 869;
Best Local Similarity 46.9%; Pred. No. 1;
Matches 135; Conservative 0; Mismatches 153; Indels 0; Gaps 0;

QY 632 GGCCCTGACGGTGCTTCAGGTGCTCCGCTGTCCTGTTCTTGGTATTACGAGCGCTGGATCTTC 591
Db 330 GTTCCTTGCTGTTGGTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 389

QY 692 GGTGTGAACCTTCGCTCGCTGGCGGTTCTTCATCGGTATCGGTACCGCAGCTATCGGT 751
Db 390 GTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 449

QY 752 GCGCGTTTCATTGCACTCTTTAATATCAAGGAGTTGCGTTGGCGCTGCAGGTTTCTTG 811
Db 450 GTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 509

QY 812 GGTGTTGTTTCTATTGATGCTCCAGATATGCTCACTTCTTGGTGTGTCAGTTGTTACC 871
Db 510 GTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 569

QY 872 TTCCTTCATCGCATTCGGCGCAGGATTCCTTATGGCCTTACTTGGTT 919
Db 570 GTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 617

```

Search completed: October 30, 2004, 17:04:56
Job time : 4442 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 29, 2004, 22:57:53 ; Search time 197 Seconds

(without alignments)

1366.880 Million cell updates/sec

Title: US-09-604-231-2

Perfect score: 2363

Sequence: 1 MAMVFLVNGYDVATMAA.....IEAGANLLNVAKEAVPATP 468

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 02.*

1: uniprot_prot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2363	100.0	661	2	Q8NMD6
2	2363	100.0	661	2	CAF21304
3	1062.5	45.0	651	2	Q88ZV9
4	1058.5	44.8	651	1	PTSA_PDPDE
5	977	41.3	534	2	Q93ML1
6	962	40.7	650	2	Q9S6S6
7	953.5	40.4	664	1	PTSA_STRMU
8	914.5	38.7	647	2	Q74HI8
9	914.5	38.7	647	2	AAS09702
10	901	38.1	649	2	Q7WNP7
11	832.5	35.2	643	2	Q82YR5
12	759.5	32.1	627	2	Q9L8G6
13	715	30.3	626	2	Q831B4
14	701	29.7	480	1	PTSB_STAXY
15	695.5	29.4	630	2	Q8KGI9
16	695	29.4	617	2	Q8Y904
17	678	28.7	474	1	PTSB_PASMU
18	674	28.5	633	2	Q82FS7
19	673.5	28.5	636	1	PTBA_BACHD
20	673	28.5	655	2	Q97NW9
21	670.5	28.4	644	2	Q8DUF2
22	667	28.2	634	2	Q8YAT6
23	667	28.2	634	2	Q725E2
24	667	28.2	634	2	AAT02819
25	664	28.1	481	2	Q8CN82
26	664	28.1	627	2	Q97PB8
27	663.5	28.1	676	2	Q8E208
28	660	27.9	627	2	Q8DNS8
29	659	27.9	479	2	Q9KLT8
30	659	27.9	480	2	Q8G6U2
31	659	27.9	480	2	Q8NV35

32	659	27.9	480	2	Q99RQ0
33	659	27.9	480	2	Q7A3V6
34	659	27.9	480	2	Q88SA6
35	658.5	27.9	644	2	Q9KJ80
36	658	27.8	480	2	Q6GE54
37	658	27.8	705	2	Q8DN16
38	657	27.8	609	1	PTBA_BACSU
39	656.5	27.8	676	2	Q8E7G5
40	655	27.7	655	2	Q8KST1
41	654.5	27.7	674	2	Q79W20
42	654.5	27.7	674	2	Q8K5K9
43	652	27.6	627	2	Q8NZJ7
44	648.5	27.4	624	2	Q89SB9
45	648.5	27.4	674	2	Q8NZ40

ALIGNMENTS

RESULT 1

Q8NMD6 PRELIMINARY; PRT; 661 AA.
ID Q8NMD6 Q6M2J8;
AC Q8NMD6; Q6M2J8;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Phosphotransferase system IIC components, glucose/maltose/N-acetylglucosamine-specific (EC 2.7.1.69) (ENZYME II SUCROSE PROTEIN).
DE GN Name=ptsS; OrderedLocNames=Cgl2642, c92925;
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxId=1718;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025 Nakagawa;
RA Nakagawa S.;
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RX PubMed=12948626;
RA Kalinowski J., Bathe B., Bartels D., Bischoff N., Bott M., Burkovski A., Dusch N., Eggeling L., Eikmanns B.J., Gaigalat L., Goesmann A., Hartmann M., Huthmacher K., Kraemer R., Linke B., McHardy A.C., Meyer F., Moeckel B., Pfeifferle W., Puhler A., Rey D.A., Rueckert C., Rupp O., Sahn H., Wendisch V.F., Wiegand I., Tauch A.;
RT "The complete Corynebacterium glutamicum ATCC 13032 genome sequence and its impact on the production of L-aspartate-derived amino acids and vitamins.";
RL EMBL; AF005282; BAC00036.1; -.
DR EMBL; BX927155; CAF21304.1; -.
DR HSP; P45618; 2GPR.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008982; F:protein-N(P)-phosphohistidine-sugar phospho. .; IEA.
DR GO; GO:0005351; F:sugar porter activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0009401; P:phosphoenolpyruvate-dependent sugar phospho. .; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR011055; Dup_hybrid_motif.
DR InterPro; IPR001996; Ptrans_EIIB.
DR InterPro; IPR003352; Ptrans_EIIC.
DR InterPro; IPR01127; Ptrans_EIIA.
DR InterPro; IPR010973; Ptrans_EIIB_sucr.
DR Pfam; PF00358; Ptrans_EIIA_1; I.
DR Pfam; PF00367; Ptrans_EIIB; I.
DR Pfam; PF02378; Ptrans_EIIC; I.
DR ProDom; PD001476; Ptrans_EIIB; I.
DR ProDom; PD002243; Ptrans_EIIA; I.

DR TIGRFAMS; TIGR00830; PTB; 1.
DR TIGRFAMS; TIGR01996; PTS-II-BC-sucr; 1.
DR PROSITE; PS01035; PTS E11B CYS; 1.
KW Complete proteome; Transferase.
SQ SEQUENCE 661 AA; 69148 MW; 69148 MW; COE5E1C7833B9F74 CRC64;

Query Match 100.0%; Score 2363; DB 2; Length 661;
Best Local Similarity 100.0%; Pred. No. 3.6e-144;
Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAMVFPVLNGYDVAAATMAAGEMPMNSLFLGLDVAQAGYQGTGVLVPLVSVILATIERFLH 60
DB 194 MAMVFPVLNGYDVAAATMAAGEMPMNSLFLGLDVAQAGYQGTGVLVPLVSVILATIERFLH 253
QY 61 KRLKGTADFLITPVLITLLTGFLTFIAIGPMRWVGVDLAHLGQLGDFGPGVGLLFLGL 120
DB 254 KRLKGTADFLITPVLITLLTGFLTFIAIGPMRWVGVDLAHLGQLGDFGPGVGLLFLGL 313
QY 121 VYSPVITGLHQSPPPIELFELFNQGGSFIFATASMANIAQGAACLAFFFLAKSEKLGKLA 180
DB 314 VYSPVITGLHQSPPPIELFELFNQGGSFIFATASMANIAQGAACLAFFFLAKSEKLGKLA 373
QY 181 GASGVSALVIGTEPAIFGVNLRLEWPFIFIGTAAIGGALIALFNKIKAVAGAAAGFLGV 240
DB 374 GASGVSALVIGTEPAIFGVNLRLEWPFIFIGTAAIGGALIALFNKIKAVAGAAAGFLGV 433
QY 241 SIDAPDMVMFLVCAVVTFFIAFGAAIAYGLVLRNGSIDPDATAAPVPGTTKAEABAP 300
DB 434 SIDAPDMVMFLVCAVVTFFIAFGAAIAYGLVLRNGSIDPDATAAPVPGTTKAEABAP 493
QY 301 AEFNSDSTIIQAPLTGEAIALSSVSDAMFASGKLGSGVAIVPTKQLVSPVSGKIIVAFP 360
DB 494 AEFNSDSTIIQAPLTGEAIALSSVSDAMFASGKLGSGVAIVPTKQLVSPVSGKIIVAFP 553
QY 361 SCHFAVTKAEDGNSVDILMHIGFTVNLNGTHFNPLKQGDVEKAGELLCFDDAIAK 420
DB 554 SCHFAVTKAEDGNSVDILMHIGFTVNLNGTHFNPLKQGDVEKAGELLCFDDAIAK 613
QY 421 AAGYEVTTPVIVSNYKKTGPVNTYGLGEIEAGANLLNVAKKEAVPATP 468
DB 614 AAGYEVTTPVIVSNYKKTGPVNTYGLGEIEAGANLLNVAKKEAVPATP 661

RESULT 3
Q88ZV9 PRELIMINARY; PRT; 651 AA.
ID Q88ZV9; PRELIMINARY; PRT; 651 AA.
AC Q88ZV9; PRELIMINARY; PRT; 651 AA.
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Sucrose PTS, E11B (EC 2.7.1.69).
GN Name=pts1BCA; OrderedLocusNames=lp_0185;
OS Lactobacillus plantarum.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=1590;
RN [1]_TaxID=1590;
RP SEQUENCE FROM N.A.
RC STRAIN=NCIMB 8826 / WCFS1;
RX MEDLINE=22480296; PubMed=12566566;
RA Kleerebezem M., Boekhorst J., van Kranenburg R., Molenaar D.,
RA Kuipers O.P., Leer R., Tarchini R., Peters S.A., Sandbrink H.M.,
RA Fiers M.W.E.J., Stiekema W., Klein Lankhorst R.M., Bron P.A.,
RA Hoffer S.M., Nierop Groot M.N., Kerkhoven R., De Vries M., Ursing B.,
RA De Vos W.M., Siezen R.J.;
RA "Complete genome sequence of Lactobacillus plantarum WCFS1.";
RT Proc. Natl. Acad. Sci. U.S.A. 100:1990-1995(2003).
RL EMBL; AL935252; CAD62855.1; -.
DR HSSP; P20166; 1GPR.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008982; F:sugar porter activity; IEA.
DR GO; GO:0005351; F:sugar porter activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0009401; P:phosphoenolpyruvate-dependent sugar phospho. .; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR011055; Dup.hybrid motif.
DR InterPro; IPR001996; Ptrans_E11B.
DR InterPro; IPR003352; Ptrans_E11C.

DR TIGRFAMS; TIGR00830; PTB; 1.
DR TIGRFAMS; TIGR01996; PTS-II-BC-sucr; 1.
DR PROSITE; PS01035; PTS E11B CYS; 1.
KW Complete proteome; Transferase.
SQ SEQUENCE 661 AA; 69148 MW; 69148 MW; COE5E1C7833B9F74 CRC64;

Query Match 100.0%; Score 2363; DB 2; Length 661;
Best Local Similarity 100.0%; Pred. No. 3.6e-144;
Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAMVFPVLNGYDVAAATMAAGEMPMNSLFLGLDVAQAGYQGTGVLVPLVSVILATIERFLH 60
DB 194 MAMVFPVLNGYDVAAATMAAGEMPMNSLFLGLDVAQAGYQGTGVLVPLVSVILATIERFLH 253
QY 61 KRLKGTADFLITPVLITLLTGFLTFIAIGPMRWVGVDLAHLGQLGDFGPGVGLLFLGL 120
DB 254 KRLKGTADFLITPVLITLLTGFLTFIAIGPMRWVGVDLAHLGQLGDFGPGVGLLFLGL 313
QY 121 VYSPVITGLHQSPPPIELFELFNQGGSFIFATASMANIAQGAACLAFFFLAKSEKLGKLA 180
DB 314 VYSPVITGLHQSPPPIELFELFNQGGSFIFATASMANIAQGAACLAFFFLAKSEKLGKLA 373
QY 181 GASGVSALVIGTEPAIFGVNLRLEWPFIFIGTAAIGGALIALFNKIKAVAGAAAGFLGV 240
DB 374 GASGVSALVIGTEPAIFGVNLRLEWPFIFIGTAAIGGALIALFNKIKAVAGAAAGFLGV 433
QY 241 SIDAPDMVMFLVCAVVTFFIAFGAAIAYGLVLRNGSIDPDATAAPVPGTTKAEABAP 300
DB 434 SIDAPDMVMFLVCAVVTFFIAFGAAIAYGLVLRNGSIDPDATAAPVPGTTKAEABAP 493
QY 301 AEFNSDSTIIQAPLTGEAIALSSVSDAMFASGKLGSGVAIVPTKQLVSPVSGKIIVAFP 360
DB 494 AEFNSDSTIIQAPLTGEAIALSSVSDAMFASGKLGSGVAIVPTKQLVSPVSGKIIVAFP 553
QY 361 SCHFAVTKAEDGNSVDILMHIGFTVNLNGTHFNPLKQGDVEKAGELLCFDDAIAK 420
DB 554 SCHFAVTKAEDGNSVDILMHIGFTVNLNGTHFNPLKQGDVEKAGELLCFDDAIAK 613
QY 421 AAGYEVTTPVIVSNYKKTGPVNTYGLGEIEAGANLLNVAKKEAVPATP 468
DB 614 AAGYEVTTPVIVSNYKKTGPVNTYGLGEIEAGANLLNVAKKEAVPATP 661

RESULT 2
CAF21304 PRELIMINARY; PRT; 661 AA.
ID CAF21304; PRELIMINARY; PRT; 661 AA.
AC CAF21304; PRELIMINARY; PRT; 661 AA.
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE ENZYME II SUCROSE PROTEIN (EC 2.7.1.69).
GN PTSS OR CG2925.
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Corynebacteriidae; Corynebacterium.
OX NCBI_TaxID=17118;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RX PubMed=12948626;
RA Kalinowski J., Bathe B., Bartels D., Bischoff N., Bott M.,
RA Burkowski A., Busch N., Eggeling L., Eikmanns B.J., Gaigalat L.,
RA Goemann A., Hartmann M., Huchtmacher K., Kraemer R., Linke B.,
RA McHardy A.C., Meyer F., Moeckel B., Pfeifferle W., Puhler A.,
RA Rey D.A., Rueckert C., Rupp O., Sahn H., Wendisch V.F., Wiegand I.,
RA Tauch A.;
RA "The complete Corynebacterium glutamicum ATCC 13032 genome sequence
RT and its impact on the production of L-aspartate-derived amino acids
RT and vitamins.";
RL J. Biotechnol. 104:5-25(2003).
DR EMBL; BX927155; CAF21304.1; -.
KW Transferase.
SQ SEQUENCE 661 AA; 69148 MW; 69148 MW; COE5E1C7833B9F74 CRC64;

	InterPro; IPR001127; PTS_EIIA.	
DR	InterPro; IPR001973; PTS_II_BC_sucr.	
DR	Pfam; PF00358; PTS_EIIA.1; 1.	
DR	Pfam; PF00367; PTS_EIIB; 1.	
DR	Pfam; PF02378; PTS_EIIC; 1.	
DR	ProDom; PD001476; Etrans_EIIB; 1.	
DR	ProDom; PD002243; PTS_EIIA; 1.	
DR	TIGRFAMs; TIGR00830; PTBA; 1.	
DR	TIGRFAMs; TIGR01996; PTS-II-BC-sucr; 1.	
DR	PROSITE; PS00371; PTS_EIIA.1; 1.	
DR	PROSITE; PS01035; PTS_EIIB_CYS; 1.	
KW	Complete proteome; transference.	
SQ	SEQUENCE 651 AA; 68514 MW; CBFBS3SEDDDB73DD6 CRC64;	

	Query Match	45.0%; Score 1062.5; DB 2; Length 651;
	Best Local Similarity	46.7%; Pred. No. 2,7e-60;
	Matches 221; Conservative 80; Mismatches 149; Indels 23; Gaps 5;	

Qy	1	MAMVPPSLVNGYDVAAATMAAGEMPMSLFLGLDVAQAGYGTVLPVLIVVSUILATIEKFLH	60
Dd	194	MIMVLPSSLVNGYSVATTMAAGKMVVWNVEGHLHVAQAGYQGQVLPVLAVAYILATLEKEFH	253
Qy	61	KRLXGTADFLITPVLTLLTGLFLEFIATGPAMRWGVDLAHGLOGLYDFGGPVGGLLPLGL	120
Dd	254	XHIKGAEDFTTPFAIVITGLFTFTIVGPLRTVSDALTNGLVGLYNTTWGMGIFGL	313
Qy	121	VYSPVITGLHQSPPIELELF-----NOGSFFIFATASMANIAOGAACLAVFELAKSEKL	176
Dd	314	LYSALVITGLHOTPALETQLLANVAKTGGSFIPVASMANIGOCAATLAIFFATKSQKQ	373
Qy	177	KGLAGAGSVNAVLTGITPEAPFGVNLRWRPFFIGTAAIGCALIFLNIXAVALGAAGF	236
Dd	374	KALTSSAGVSALLGITPEAPFGVNLMKMPFVFEAAIASGIASAFLGLFHVLSVANGPASV	433
Qy	237	LGVSVIDAPDMVMFLVCNVITFTPTARGAATAGYLYVRNGSIDPDATAAPVPAGTTKAE	296
Dd	434	IGFTISASKSIPAPMLSAVISFVVAFTFTFY----AKRTLGDDRDQVKSPAPTSTV---	486
Qy	297	AEPAPFNSNSTIIQAPLTGEAIALSSVDAMFASGLGSVATVPKGLQVSPVSKIV	356
Dd	487	-----INVNDIELISAPVTGTSLSLKQNDOVFSRAEIMGKAALVPSAQOVVAPADGVIT	540
Qy	357	VAPPSGHAFVRTKABDSNVDDLHMHGDTVNLNGTHFNPKLKGQDEVKAGELLCEFDI	416
Dd	541	VTYDSHHAYGIKTWA---GAELIHGLDVLNNGEHFTTNVQKGTUVHQEDLLGTTFDV	596
Qy	417	DAIKAAQGEVTPIVLVSNYKKTGPNVTYIGETBEGANLLNVAKEEA--VPAT	467
Dd	597	AALKAAANDPTVMLVINTANYANVERLKNVTNQAGEQLVALTEPTASSVAAT	649

sugar phosphotransferase system (pts), a major carbohydrate active
-transport system. The IIDC domains contain the sugar binding site
and the transmembrane channel; the IIA domain contains the primary
phosphorylation site (the donor is phospho-HP); IIA transfers its
phosphoryl group to the IIB domain which finally transfers it to
the sugar.

!- CATALYTIC ACTIVITY: Protein N-phosphohistidine + sugar = protein
histidine + sugar phosphate.

!- SUBCELLULAR LOCATION: Integral membrane protein.

!- SIMILARITY: Contains 1 PTS EIIA domain.

!- SIMILARITY: Contains 1 PTS EIIB domain.

!- SIMILARITY: Contains 1 PTS EIIC domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@sb-sib.ch).

EMBL; Z32771; CAA83668.1; --
EMBL; L32093; AAA25567.1; --
PIR; S44257; S44257.
HSSP; P20166; 1GPR.

InterPro; IPR011055; Dup hybrid_motif.
InterPro; IPR001996; Ptrans_EIIB.
InterPro; IPR003352; Ptrans_EIIC.
InterPro; IPR001127; PTS_EIIA.
InterPro; IPR010973; PTS_II_BC_sucr.
Pfam; PF00359; PTS_EIIA_1; I.
Pfam; PF00367; PTS_EIIB; 1.
Pfam; PF02378; PTS_EIIC; 1.
ProDom; PD002243; PTS_EIIA; 1.
ProDom; PD001476; PTS_EIIB; 1.
TIGRFAMs; TIGR00826; EIIB_glc; 1.
TIGRFAMs; TIGR00830; FTBA; 1.
TIGRFAMs; TIGR01995; PTS-II-ABC-beta; 1.
TIGRFAMs; TIGR01996; PTS-II-BC-sucr; 1.
TIGRFAMs; TIGR01992; PTS-IIBC-Tre; 1.
PROSITE; PS00371; PTS_EIIA_1; 1.
PROSITE; PS01035; PTS_EIIB_CYS; 1.
Phosphorylation; Phosphotransferase system; Sugar transport;
Transferase; Transmembrane.

DOMAIN 1 40 EIIB.
DOMAIN 111 418 EIIC.
DOMAIN 510 651 EIIA.
MOD_RES 25 25 Phosphocysteine (By similarity).
MOD_RES 324 324 Phosphohistidine (By similarity).
MOD_RES 562 562 Phosphohistidine (By similarity).
SEQUENCE 651 AA; 68454 MW; C87BA09D550A77F8 CRC64;

CC -!- FUNCTION: This is a component of the phosphoenolpyruvate-dependent

Db	434	IGISTASKSIPAFMLSAVISFVFAIPFTIY-----AKRTLGDDRDVOVKSPTISIV---	486
QY	297	AEAPAFSNDSTIIQAPLTGEALIALSVSDAMFASGKLGSGVAIVPTKGOLVSPVSGKIV	356
Db	487	-----INVNDIISAPVTGASESLKQVNDQVFSAEIMGKGAALVPSSDQWAPADGVIT	540
QY	357	VAPPSGHAFVAVRTKAEDGSNDVILMHIGFTVNLNGTHFNPLKKQDGEVKAAGELLCEFDI	416
Db	541	VTDYSHAYGIKTTA-----GAEIILHGLDVTNLNGEHTTNNVQKGDVTHQGDLLGTFDI	596
QY	417	DAIKAAQVEVTPPIVVSNNKKTGPTVNTYGLGEIEAGANL--LNVAKKEAVPAT	467
Db	597	AALKAANYDPTVMLIVNTINYANVERLKVNTVQAGEQLVALTAPAASSVAAT	649
RESULT 5			
Q93ML1	ID	Q93ML1	PRELIMINARY; PRT; 534 AA.
AC	Q93ML1	01-DEC-2001 (TrEMBLrel. 19, Created)	
DT	01-DEC-2001	(TrEMBLrel. 19, Last sequence update)	
DT	01-MAR-2004	(TrEMBLrel. 26, Last annotation update)	
DE	Sucrose-specific enzyme II of the PTS (Fragment).		
GN	Name=sacr;		
OS	Lactobacillus sakei.		
OC	Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;		
OC	Lactobacillus.		
OX	NCBI_TaxID=1599;		
EN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=21822499; PubMed=11832506;		
RA	Dudez A.-M., Chailion S., Zagorec M., Stentz R.,		
RA	Championier-Verges M.C., Albert C.A., Zagorec M.;		
RT	"Physical and Genetic map of the Lactobacillus sakei 23K chromosome."		
RL	Microbiology 148:421-431(2002).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RA	Dudez A.-M., Chailion S., Zagorec M.;		
RA	Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AF401046; AAK92528.1; -.		
DR	HSSP; P45618; 2GPR.		
DR	GO; GO:0016020; C:membrane; IEA.		
DR	GO; GO:0008982; P:protein-N(P1)-phosphohistidine-sugar phospho. . . ; IEA.		
DR	GO; GO:0005351; F:sugar porter activity; IEA.		
DR	GO; GO:0009401; P:phosphoenolpyruvate-dependent sugar phospho. . . ; IEA.		
DR	GO; GO:0006810; P:transport; IEA.		
DR	InterPro; IPR011055; Dup hybrid motif.		
DR	InterPro; IPR003352; Ptrans EIIC.		
DR	InterPro; IPR001127; PTS EIIA.		
DR	Pfam; PF00358; PTS_EIIA_1; 1.		
DR	Pfam; PF02378; PTS_EIIC; 1.		
DR	TIGRFAMs; TIGR00830; PTBA; 1.		
DR	PROSITE; PS00371; PTS_EIIA_1; 1.		
FT	NON_TER		
SQ	SEQUENCE	534 AA; 56230 MW; 01804F9DE70C0089 CRC64;	
Query Match			
Best Local Similarity 41.3%; Score 977; DB 2; Length 534;			
Matches 211; Conservative 90; Mismatches 138; Indels 38; Gaps 10;			
QY	1	MAWVPSLVNGYDVAAATMAAGEMPMSLFGDLVAQAGYQGTVLPLVLSWILLATIEKFLH	60
Db	81	MLVMPNLVNGYGVASISATGHMTYWHVFLNIAQAGYQGVLPVGVAFILANLEKFFH	140
QY	61	KRLKGTADFLITPLVLTLLTGLTFTIAIGFAMWVGVDVLAHGLQGHYDFGPGVGLLFL	120
Db	141	KLHNDVDTFTFPLMSLIITGLTFTTLVGLPALIVSGVTDLSVWAYQTLGAVGMGIFGL	200
QY	121	VYSPVITGLHQSFPIELF---NQGSFTAFASMANIAQGAACLAIVFLAKSEKL	176
Db	201	GYSALVLTGLHQSFPAIETLLADIATKGTGSFFPVAAMANTIAQGAATFAVFEVTKNQ	260
RESULT 6			
Q9S6S6	ID	Q9S6S6	PRELIMINARY; PRT; 650 AA.
AC	Q9S6S6;		
DT	01-MAY-2000	(TrEMBLrel. 13, Created)	
DT	01-MAY-2000	(TrEMBLrel. 13, Last sequence update)	
DT	01-MAR-2004	(TrEMBLrel. 26, Last annotation update)	
DE	Enzyme II sucrose protein (EC 2.7.1.69).		
GN	Name=sacr;		
OS	Lactococcus lactis.		
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.		
OX	NCBI_TaxID=1358;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=99173919; PubMed=10074089;		
RA	Luesink E.J., Marugg J.D., Kuipers O.P., De Vos W.M.;		
RT	"Characterization of the divergent sacBK and sacAR operons, involved		
RT	in sucrose utilization by Lactococcus lactis."		
RL	J. Bacteriol. 181:1924-1926(1999).		
DR	EMBL; Z97015; CAB09690.1; -.		
DR	HSSP; P45618; 2GPR.		
DR	GO; GO:0016020; C:membrane; IEA.		
DR	GO; GO:0008982; P:protein-N(P1)-phosphohistidine-sugar phospho. . . ; IEA.		
DR	GO; GO:0005351; F:sugar porter activity; IEA.		
DR	GO; GO:0016740; P:transferase activity; IEA.		
DR	GO; GO:0009401; P:phosphoenolpyruvate-dependent sugar phospho. . . ; IEA.		
DR	GO; GO:0006810; P:transport; IEA.		
DR	InterPro; IPR011055; Dup hybrid motif.		
DR	InterPro; IPR001996; Ptrans EIIB.		
DR	InterPro; IPR003352; Ptrans EIIC.		
DR	InterPro; IPR001127; PTS EIIA.		
DR	InterPro; IPR010973; PTS_EIIA_1; 1.		
DR	Pfam; PF00358; PTS_EIIB; 1.		
DR	Pfam; PF02378; PTS_EIIC; 1.		
DR	Pfam; PF00367; PTS_EIIB; 1.		
DR	ProbDom; PD001476; Ptrans EIIB; 1.		
DR	ProbDom; PD002243; PTS_EIIA; 1.		
DR	TIGRFAMs; TIGR00830; PTBA; 1.		
DR	TIGRFAMs; TIGR01996; PTS-II-BC-sucr; 1.		
DR	PROSITE; PS00371; PTS_EIIA_1; 1.		
DR	PROSITE; PS01035; PTS_EIIB_CYS; 1.		
KW	Transferase.		
SQ	SEQUENCE	650 AA; 69636 MW; 250E26F7664D4204 CRC64;	
Query Match			
Best Local Similarity 40.7%; Score 962; DB 2; Length 650;			
Matches 203; Conservative 89; Mismatches 145; Indels 36; Gaps 7;			
QY	1	MAWVPSLVNGYDVAAATMAAGEMPMSLFGDLVAQAGYQGTVLPLVLSWILLATIEKFLH	60
Db	194	MLVMPNLVNGYGVASISATGHMTYWHVFLNIAQAGYQGVLPVGVAFILAKLRRFFH	253


```
QY 228 AVALGAGFGLGVSVISIDAPDMVMFLVCAVVTFFIAFGAIAIYGLYLVRNGSIDPDATAAP 287
Db 434 AVLSGAGFGLGFSIKASSIPFYVWCLEISFAIAFAVITYGYG-----KTKAVDVFAAEAA 488
QY 288 VPAGTTKABEAEAPAEFNSDST-----IIOAPLTGEAIALSSVSDFAMFASGKLGSVAI 340
Db 489 VERAIEVQ-EIPEEAASAANKAQVTDVLEAAPLAGAEVLTSDVPFSEANGKGIAT 547
QY 341 VPTKGQIVSPVSGKIVVAFPSGHAFVRTKADGNSVDIIMHIGFTDVNLNGTHFNPLK 400
Db 548 KPSGNTVYAPVDGTQVIAFTGHAYGI--KSDNGA--EILIHIGIDTVSMEGKGFQKVQ 603
QY 401 QGDEVKAGELLCEPDIDAIKAAGYEVTTPVVSNNYKKTGPWNTY-GLGEIEAGANLINVA 459
Db 604 ADQKIKKGDVLGTFDSKIAEAGLDNTTWFIVTNTADYASVETLASSGTVAVGDSLLEVK 663
QY 460 K 460
Db 664 K 664

RESULT 8
Q74HI8
ID Q74HI8 PRELIMINARY; PRT; 647 AA.
AC Q74HI8;
DT 05-JUL-2004 (TremBLrel. 27, Created)
DT 05-JUL-2004 (TremBLrel. 27, Last sequence update)
DE Phosphoenolpyruvate-dependent sugar phosphotransferase system EIIABC,
DE sucrose specific.
GN OrderedLocusNames=J0519;
OS Lactobacillus johnsonii.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=33959;
RN [1]_TaxID=33959;
RP SEQUENCE FROM N.A.
RC STRAIN=NCC 533;
RX PubMed=14966310;
RA Pridmore R.D., Berger B., Desiere F., Vilanova D., Barretto C.,
RA Pittet A.-C., Zwaalen M.-C., Rouvet M., Altermann E., Barrangou R.,
RA Mollet B., Mercenier A., Klaenhammer T., Arigoni F., Schell M.A.;
RT "The genome sequence of the probiotic intestinal bacterium
RT Lactobacillus johnsonii NCC 533.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:2512-2517(2004).
DR EMBL; AE017206; AAS09702.1; -.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR011055; Dup hybrid motif.
DR InterPro; IPR011535; EIIB_glc.
DR InterPro; IPR001996; Ptrans_EIIB.
DR InterPro; IPR003352; Ptrans_EIIC.
DR InterPro; IPR011127; PTS_EIIA.
DR InterPro; IPR010973; PTS_II_BC_sucr.
DR Pfam; PF00358; PTS_EIIA_1; 1.
DR Pfam; PF00367; PTS_EIIB; 1.
DR Pfam; PF02378; PTS_EIIC; 1.
DR ProDom; PD001476; Ptrans_EIIB; 1.
DR ProDom; PD002243; PTS_EIIA; 1.
DR TIGRFAMS; TIGR00826; EIIB_glc; 1.
DR TIGRFAMS; TIGR00830; PTBA; 1.
DR TIGRFAMS; TIGR01996; PTS-II-BC-sucr; 1.
DR PROSITE; PS00371; PTS_EIIA_1; 1.
DR PROSITE; PS01035; PTS_EIIB_CUS; 1.
DR Complete proteome; Pyruvate; Transferrase.
SQ SEQUENCE 647 AA; 69202 MW; 95FA74C533CE9541 CRC64;

Query Match 38.7%; Score 914.5; DB 2; Length 647;
Best Local Similarity 41.3%; Pred. No. 9.5e-51;
Matches 195; Conservative 88; Mismatches 160; Indels 29; Gaps 6;

QY 1 MANVFPPLVNGYDVAAATMAAGEMPWMSLFGLDVAQAGYQGTVLPVLVSVWILATIEKFLH 60
Db 194 MIMTTPAL-----GGAGKYWDIFGLHVSQTNVQYQVLPVLVAVWLLAFLEKRFH 242
QY 61 KRLKGTADFLITPVLTLLTGLTFTFAIGPAMRWGVDVLAHGLQGLYDFGPGVGLLFLGL 120
Db 243 KWLPSAVDTFTPLLSIMITGFTFTIIGVFKGVSDAITAAIWLNTTGAFCMGIFGL 302
QY 121 VYSPVITGLHQSFPPIELEFNQ-----GGSFIFATASMANIAQGAACLAFFFLAKSE 174
Db 303 SYSAIVTTGLHQSFPVAVETQLLAEYARGRSGDFIFVTACMANVAQGAATFAIYFLTKNE 362
QY 175 KLKGLGASGVSAVLGITEPAIFGNVLRRLRPPIGIGTAAIGGALIALFNKIKAVALLGAA 234
Db 363 KVKGLASSGVSALLGITEPALFGVNLKYKFPFPCALIGVGAAPAGLHVHTAALGSA 422
QY 235 GFLGVSIDAPDMVMFLVCAVVTFFIAFGAIAIYGLY-----LVRRNGSIDPDATAAPVP 289
Db 423 GFLGFLSMVFKSIPMWALSIVSIFVAFGLTFVYGRHKFKEDVVEESGTVF---SAGDQV 479
QY 290 AGTTKABEAPAEFNSDSTIIIOAPLTGEAIALSSVSDFAMFASGKLGSVAIVPTKGQIVS 349
Db 480 AQEKARQIILEDKELHDEVIAPVSGKAEBSLKDNDVDFSTFSTANGKGAANVPSEGTIYS 539
QY 350 PVSGKIVVAFPSGHAFVRTKADGNSVDIIMHIGFTDVNLNGTHFNPLKQSGDEVKAGE 409
Db 540 PVTGEITVAVETKHAYGI--KSDEGA--EVLIHIGIDTVNMKGEGFTTDVKQGGHVEKGE 595
QY 410 LLCEFDIDAIIKAAGYEVTTPVVSNNYKKTGPWNTYGLGEIEAGANLINVAKK 461
Db 596 KLGTVDDLDAVKKAGYDITVMTVIINTTSYANVQRIDGVEKKHGDLLAVTKR 647

RESULT 9
AAS09702
ID AAS09702 PRELIMINARY; PRT; 647 AA.
AC AAS09702;
DT 02-MAR-2004 (TremBLrel. 27, Created)
DT 02-MAR-2004 (TremBLrel. 27, Last sequence update)
DE Phosphoenolpyruvate-dependent sugar phosphotransferase system EIIABC,
DE sucrose specific.
GN LJ0519
OS Lactobacillus johnsonii.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=33959;
RN [1]_TaxID=33959;
RP SEQUENCE FROM N.A.
RC STRAIN=NCC 533;
RX PubMed=14966310;
RA Pridmore R.D., Berger B., Desiere F., Vilanova D., Barretto C.,
RA Pittet A.-C., Zwaalen M.-C., Rouvet M., Altermann E., Barrangou R.,
RA Mollet B., Mercenier A., Klaenhammer T., Arigoni F., Schell M.A.;
RT "The genome sequence of the probiotic intestinal bacterium
RT Lactobacillus johnsonii NCC 533.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:2512-2517(2004).
DR EMBL; AE017206; AAS09702.1; -.
DR Pyruvate; Transferrase.
SQ SEQUENCE 647 AA; 69202 MW; 95FA74C533CE9541 CRC64;

Query Match 38.7%; Score 914.5; DB 2; Length 647;
Best Local Similarity 41.3%; Pred. No. 9.5e-51;
Matches 195; Conservative 88; Mismatches 160; Indels 29; Gaps 6;

QY 1 MANVFPPLVNGYDVAAATMAAGEMPWMSLFGLDVAQAGYQGTVLPVLVSVWILATIEKFLH 60
Db 194 MIMTTPAL-----GGAGKYWDIFGLHVSQTNVQYQVLPVLVAVWLLAFLEKRFH 242
QY 61 KRLKGTADFLITPVLTLLTGLTFTFAIGPAMRWGVDVLAHGLQGLYDFGPGVGLLFLGL 120
Db 243 KWLPSAVDTFTPLLSIMITGFTFTIIGVFKGVSDAITAAIWLNTTGAFCMGIFGL 302
QY 121 VYSPVITGLHQSFPPIELEFNQ-----GGSFIFATASMANIAQGAACLAFFFLAKSE 174
```

```

Db 303 SYSAIVTTGLHQSFPFAVETOLLAEYARGSGDGFIVTACMANVAQGAATFAIYFLTKNE 362
QY 175 KIKGLAGASGVSAVLGITEPAIGFVNLRLWPPFFIGITAAIGALIALFNKKAVALGAA 234
Db 363 KVKGLASSGVSAVLLGITEPAIGFVNLRLWPPFFIGITAAIGALIALFNKKAVALGAA 422
QY 235 GFLGVVSDIDAPDMVFLVCVVVTFPIAFGAIAIYGLY-----LVRNGSIDDPDATAAPVP 289
Db 423 GFLGVVSDIDAPDMVFLVCVVVTFPIAFGAIAIYGLY-----LVRNGSIDDPDATAAPVP 289
QY 290 AGTTCARAEAPAEFSDNSTIIQAPLTGEATASVSDAMFASGKLGSGVAIVTKGOLVS 349
Db 480 AQEKAEQIIEKDELHDEVAIPVSGAESLKDNDVDFESTRANGGAAMVSEGTIYS 539
QY 350 PVSGKIVVAPPSGHAFVTRKABDGSNDVILMHIGFDTVNLNGTHFNPLKKGQDEVKAGE 409
Db 540 PVTGEITVAYETKHAYGI--KSDEGA--EVLHIGDITVNMKGEGFTYDKQQQHWKEGE 595
QY 410 LLCEFDIDAKAGYEVTTPVIVSNYKKGTPVNTYGLGEIEAGANLNAVKK 461
Db 596 KLGTVDLDAVKKAGYDTTMTVTITNTTSYANVQRIDGVEKKHGDLLIAVTKR 647

RESULT 10
Q7WNP7 PRELIMINARY; PRT; 649 AA.
AC Q7WNP7;
DT 01-OCT-2003 (TREMBLrel. 25, Created)
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
DE 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Sucrose PTS transporter (EC 2.7.1.69).
GN Name=scrA;
OS Lactobacillus acidophilus.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=1579;
RN [1]_TaxID=1579;
RP SEQUENCE FROM N.A.
RC STRAIN=NCFM;
RX PubMed=12847288;
RA Barrangou R., Altermann E., Hutkins R., Cano R., Klaenhammer T.R.;
RT "Functional and comparative genomic analyses of an operon involved in
fructooligosaccharide utilization by Lactobacillus acidophilus.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:8957-8962(2003).
DR EMBL; AY177419; AAC38866.1; -
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008982; F:protein-N(P)-phosphohistidine-sugar phospho. . .; IEA.
DR GO; GO:0005351; F:sugar porter activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0009401; F:phosphoenolpyruvate-dependent sugar phospho. . .; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR001996; P:trans_EIIB.
DR InterPro; IPR003352; P:trans_EIIC.
DR InterPro; IPR001127; PTS_EIIA.
DR InterPro; IPR010973; PTS_EIIB.
DR Pfam; PF00358; PTS_EIIA_1; 1.
DR Pfam; PF00367; PTS_EIIB; 1.
DR Pfam; PF02378; PTS_EIIC; 1.
DR ProDom; PD001476; P:trans_EIIB; 1.
DR ProDom; PD002243; PTS_EIIA; 1.
DR TIGRFAMs; TIGR00830; P:trans_EIIB; 1.
DR TIGRFAMs; TIGR01996; P:trans_EIIB; 1.
DR PROSITE; PS00371; PTS_EIIA_1; 1.
DR PROSITE; PS01035; PTS_EIIB_CYS; 1.
DR Transferrase.
SQ SEQUENCE 649 AA; 69446 MW; B4377CA24D00507B CRC64;

Query Match 38.1%; Score 901; DB 2; Length 649;
Best Local Similarity 40.8%; Pred. No. 7.1e-50;
Matches 183; Conservative 81; Mismatches 149; Indels 36; Gaps 6;

QY 1 MAMVFPVINGVDVAATWAAGEMPWWSLFGLDVAQAGQGVTLVPLVSVWILATIEKPLH 60

```

```

Db 194 MIMTTPGL-----CGATKFWDFI FGLHVAQTNQYQVIVPLVAVVWLSIFEKVFH 242
QY 61 KRLKGTADFLITVLTLLTGLTFTFAIGPAMRWGVDLAHLGQLGYDFGPGVGLLFLGL 120
Db 243 KHLPSAVDFTTFTLLSIMITGFTFTTIGPVFKGVSDAITNAIWLVDYTTGAPGMGVFGL 302
QY 121 VYSPVITGLHQSFPPIELELF-----NQGSFIFATASMANIAGAACLAFFFLAKS 173
Db 303 SYSAIVTTGLHQSFPFAVETOLLAEAFKPNPASSGDFIVTACMANVAQGAATFAIYFLTKN 362
QY 174 EKLKGLAGASGVSAVLGITEPAIGFVNLRLWPPFFIGITAAIGALIALFNKKAVALGAA 233
Db 363 KMKGLASSGVSAVLLGITEPAIGFVNLRLWPPFFIGITAAIGALIALFNKKAVALGAA 422
QY 234 AGPLGVVSDIDAPDMVFLVCVVVTFPIAFGAIAIYGLYLVVRNGSIDDPDATAAPVPAG-- 291
Db 423 AGPLGVVSDIDAPDMVFLVCVVVTFPIAFGAIAIYGLYLVVRNGSIDDPDATAAPVPAG-- 291
QY 292 -----TTKAEAEAPAEFSDNSTIIQAPLTGEATASVSDAMFASGKLGSGVAIVTKG 345
Db 477 TDYAEITQAEKIGKEQLAKDBEIIYSPVDGTPESLTKNVDQVFSAKLMGDGAIVPSDG 536
QY 346 QLVSPVSGKIVVAPPSGHAFVTRKABDGSNDVILMHIGFDTVNLNGTHFNPLKKGQDEV 405
Db 537 TIVAPVTGTVIAYETKHAYGI--KSDDGA--EVLHIGDITVNLKGEHPESVFKQGRV 592
QY 406 KAGELLCEFDIDAKAGYEVTTPVIVSN 434
Db 593 EKGDKLGSDVLDVAVKKAGYDTTMTVTITNTTSYANVQRIDGVEKKHGDLLIAVTKR 647

RESULT 11
Q82YR5 PRELIMINARY; PRT; 643 AA.
AC Q82YR5;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DE 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE PTS system, IABC components.
GN OrderedLocNames=EFA0067;
OS Enterococcus faecalis (Streptococcus faecalis).
OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX NCBI_TaxID=1351;
RN [1]_TaxID=1351;
RP SEQUENCE FROM N.A.
RC STRAIN=V583 / ATCC 700802;
RX MEDLINE=22550857; PubMed=12663927; DOI=10.1126/science.1080613;
RA Paulsen I.T., Banerjee L., Myers G.S.A., Nelson K.E., Seshadri R.,
RA Read T.D., Fouts D.E., Eisen J.A., Gill S.R., Heidelberg J.F.,
RA Tettelin H., Dodson R.J., Umayam L.A., Brinkac L.M., Beanan M.J.,
RA Dougherty S.C., DeBoy R.T., Durkin S.A., Kolonay J.F., Madupu R.,
RA Nelson W.C., Vamathevan J.J., Tran B., Upton J., Hansen T., Shetty J.,
RA Khouri H.M., Utterback T.R., Radune D., Ketchum K.A., Dougherty B.A.,
RA Fraser C.M.;
RT "Role of mobile DNA in the evolution of vancomycin-resistant
Enterococcus faecalis.";
RL Science 299:2071-2074(2003).
DR EMBL; AB016833; AAC83060.1; -
DR HSSP; P20166; IGPR.
DR TIGR; EFA0067; -
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008982; F:protein-N(P)-phosphohistidine-sugar phospho. . .; IEA.
DR GO; GO:0005351; F:sugar porter activity; IEA.
DR GO; GO:0009401; F:phosphoenolpyruvate-dependent sugar phospho. . .; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR011055; Dup hybrid motif.
DR InterPro; IPR001996; P:trans_EIIB.
DR InterPro; IPR003352; P:trans_EIIC.
DR InterPro; IPR001127; PTS_EIIA.
DR InterPro; IPR010973; PTS_EIIB.
DR Pfam; PF00358; PTS_EIIA_1; 1.
DR Pfam; PF00367; PTS_EIIB; 1.

```


DR	Pfam: PF02378; PTS_EIIC; 1.	
DR	ProDom: PD001476; Ptrans_EIIB; 1.	
DR	ProDom: PD002243; PTS_EIIA; 1.	
DR	TIGRFAMS: TIGR00830; PTBA; 1.	
DR	TIGRFAMS: TIGR01996; PTS-II-BC-sucr; 1.	
DR	PROSITE: PS00371; PTS_EIIA 1; 1.	
DR	PROSITE: PS01035; PTS_EIIB_Cys; 1.	
KW	Complete proteome; Plasmid	
SEQ	SEQUENCE 643 AA; 68532 MW; 752F65D8154AA06C CRC64;	
Query Match	35.2%; Score 832.5; DB 2; Length 643;	
Best Local Similarity	39.3%; Pred. No. 1.8e-45;	
Matches	184; Conservative 84; Mismatches 171; Indels 29; Gaps 7;	
Qy	1 MAMVFPSELVNGYDYVAATWAAGEMPWSLFGLDVAQAGYQGCVLPLVAVVSWILATIEKFLH 60	
Db	194 MIKMTPLD-----GGATRYWNIFGYHVAQTNNYAVQVLPVASVYLLSLIEKTFH 242	
Qy	61 KRLKGTADFLLTPVLTLLLTGLFTFIALGPAMRWGDVLAHGLGQLGYDFGGVPGGLLFLG 120	
Db	243 KLFSSIDFTFPLLSVIITGLFTVIGPTMLLLNSGITDAIVWLYNATGFIWGIFFG 302	
Qy	121 VYSPVITGLHQSPPIELELFNQ-----GGSFIFATASMANTAQGAACLAIVFLAKSE 174	
Db	303 TYSLIWVTGLHQSPPAIETQLLSAWTNGIGHGDFTFVVASMANVAQGAATFAIWFLTKNS 362	
Qy	175 KLGAGAGSVSAVIGITEPAIFGVNLRWPFFETIGTAAGIGALIALFNKIKAVAGAA 234	
Db	363 KTKLASAGSALIGITEPALFGVNLKRPFFCALIGSGIAAAITGLLKVAVSLGSA 422	
Qy	235 GFLGVSTIDAPMVNMFVCAVPTFFIAPGAATAYGLVLRNRNGSIDDPATAAPVAGTTK 294	
Db	423 GFLGFLSINATSIPIYLICELISFSVTALTIFYG--RTRSSSIFAERAIABQTSVDTSE 480	
Qy	295 AEAEAPAFSNDSTI--IQAPITGRIALSSVSDAMFASGKLGSGVAIVPTKGQLVSPVS 352	
Db	481 INTNQIANADEPTTIVETIVSPLAGEITITIGSVNDPVFSSESIGKGIKPNGTNTIYSPVD 540	
Qy	353 GKIVVAFPSCHAFATVTKAEDGSNVDILMHIGFDVNLNGTHFNPLKKQG--DEVKAGEL 410	
Db	541 GIVQVVFETHAYDLKS---NTGAEILTHVGIDIVSLNGKGT--KKVGAQKQVKKGEV 594	
Qy	411 LCEPFDIDAKAAGYEVVTPVIVSVNKKTPGVNTYGLGTEIAGANLINV 458	
Db	595 LGTFDSVTINSGLDTTWIVVINSKDYSEVIPKTNIVTEGAALITI 642	
RESULT 12		
Q91LG6	PRELIMINARY; PRT; 627 AA.	
ID	Q91LG6	
AC	01-OCT-2000 (TREMELrel. 15, Created)	
DT	01-OCT-2000 (TREMELrel. 15, Last sequence update)	
DT	01-OCT-2004 (TREMELrel. 28, Last annotation update)	
DE	Scra (Fusion: PTS system, beta-glucosides specific IIABC component).	
DE	Name=scra; OrderedlocusNames=CAC0423;	
OS	Clostridium acetobutylicum.	
OC	Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;	
OC	Clostridium.	
OX	NCBI_TaxID=1488;	
RN	[1]_TaxID=1488;	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=ATCC 824;	
RX	MEDLINE=20391269; PubMed=10937490;	
RA	Tangney M., Mitchell W.J.;	
RT	"Analysis of a catabolic operon for sucrose transport and metabolism	
RT	in Clostridium acetobutylicum ATCC 824."	
RL	J. Mol. Microbiol. Biotechnol. 2:71-80(2000).	
RN	[2]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=ATCC 824 / DSM 792 / VKM B-1787;	
RX	MEDLINE=21359325; PubMed=11466286;	

RESULT 13

DOI=10.1128/JB.183.16.4823-4838.2001;

RA Noelling J., Breton G., Onelchenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RA Tatusov R.J., Sabathe F., Koonette-Stamm L.A., Soucaille P.,
RA Daly M.J., Bennett G.N., Doucin E.V., Smith D.R.,
RA "Genome sequence and comparative analysis of the solvent-producing
RT bacterium *Glostridium acetobutylicum*,"
RL J. Bacteriol. 183:4823-4838(2001).
DR EMBL; AF205034; AAF38839.1; -
DR EMBL; AE007557; AAK78403.1; -
DR PIR; H96951; H96951.
DR HSP; P20166; IGPR.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008982; F:protein-N(P1)-phosphochistidine-sugar phospho. .; IEA.
DR GO; GO:0005351; F:sugar porter activity; IEA.
DR GO; GO:0009401; P:phosphoenolpyruvate-dependent sugar phospho. .; IEA.
DR GO; GO:0006810; P:transpor; IEA.
DR InterPro; IPR011055; Dup hybrid motif.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR001996; P:trans EIIB.
DR InterPro; IPR003352; P:trans EIIC.
DR InterPro; IPR001127; P:trans EIIA.
DR InterPro; IPR010973; P:trans II-BC-sucr.
DR Pfam; PF00358; P:trans EIIA_1; I.
DR Pfam; PF00367; P:trans EIIB; I.
DR Pfam; PF02378; P:trans EIIC; I.
DR ProDom; PD001476; P:trans EIIB; I.
DR ProDom; PD002243; P:trans EIIA; I.
DR TIGRFAMS; TIGR00830; P:trans EIIB; I.
DR TIGRFAMS; TIGR01996; P:trans II-BC-sucr; I.
DR PROSITE; PS00371; P:trans EIIA_1; I.
DR PROSITE; PS00134; TRYPsin_HIS; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 627 AA; 67596 MW; B835AB0238FCA436 CRC64;

Query Match 32.1%; Score 759.5; DB 2; Length 627;
Best Local Similarity 35.7%; Pred. No. 9.2e-41;
Matches 163; Conservative 80; Mismatches 176; Indels 25; Gaps 5

QY 3 MVFSLVNGYDVAAATMAAGEMPMSLFGDLVQAQGYQGTVLPLVLSVILATIEKFLHKR 62
DB 186 MIHEDLNAM---TLGEGIKHTINIFGLNIGMVGYQGTVLPLILISVWMSYIEKGLRK 241
QY 63 LKGTADFLITPLVILLTGTFTTAIGPAMRWGDLVAHGLOGLYDFGSPVGLLFLGLV 122
DB 242 VPEALDILLTFLTMITGTFFAMVVGPGFVGDEISLGLQTLTNTTTFPSGVFLFGGLY 301
QY 123 SPIVITGLHQSPFPIELELEFNQGG---SFIFATASMANIAQAAACIAYFFLAKSEKLKGL 179
DB 302 SLIVITGHHSFHAIEAGLLANPAHKNFLLPIWSMANVQAQGAALAVYFKTRDKWKSI 361
QY 180 AGASGVSAVLGITPEIPAIFGVNLRURPFFIGTGTAAIGALIALFNKVALCAAGFLGV 239
DB 362 AAPASFCLLGITPEIPAIFGVNLRVTKPTIAGALGAIGGGYIVFTKVMAMTAVGVGTIPGI 421
QY 240 VSIDAPDMVMTLVCAVVTFFIAFGAAIAYGLYLVRNGSIDPDATAAPVPAGTTRKAEABA 299
DB 422 AIVKQGSFNLIIAMI-----LAPGAGFIAMVLGIKEITTEEDLNKETVYKDIKVEEVES 477
QY 300 PAEFSNDSTIIQAPLTGBAIALSSVSDAMFASGKIGSGVAIVPTTQGLVSPVSGKIVVAF 359
DB 478 -----VVSPVNGKVLIIKNVPDKTFAEGLIGDGDGVDPGEVVSPIDGTVVHVH 527
QY 360 PSGHAFAYRTKAEDSGNVDIILMHGIFDTVNLNGTHFNPLKKGDDVKACGELCEFDIAI 419
DB 528 ETKHAIAWRKS-----NGVEMLLIHGIDIVYKMEGNGFKSFINDGEVYKGDKLIQFDLDLV 583
QY 420 KAAGYEVTTPVIVSNYKKTGPVNT 443
DB 584 KEKAVSPVLLTIVTNHEDMGFVNS 607

RESULT 13

Q831B4
ID Q831B4 PRELIMINARY; PRT; 626 AA.
AC Q831B4
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PTS system, beta-glucoside-specific IIABC component.
GN OrderedLocuNames=EF2598;
OS Enterococcus faecalis (Streptococcus faecalis).
OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX NCBI_TaxID=1351;
EN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=V583 / ATCC 700802;
RX MEDLINE=22550857; PubMed=12663927; DOI=10.1126/science.1080613;
RA Paulsen I.T., Banerjee L., Myers G.S.A., Nelson K.E., Seshadri R.,
RA Read T.D., Fouts D.E., Eisen J.A., Gill S.R., Heidelberg J.F.,
RA Tettelin H., Dodson R.J., Unayam L.A., Brinkac L.M., Beanan M.J.,
RA Daugherty S.C., Deboy R.T., Durkin S.A., Kolonay J.F., Madupu R.,
RA Nelson W.C., Vamathevan J.J., Tran B., Upton J., Hansen T., Shetty J.,
RA Khouri H.M., Uutterback T.R., Radune D., Ketchum K.A., Dougherty B.A.,
RA Fraser C.M.;
RT "Role of mobile DNA in the evolution of vancomycin-resistant
RT Enterococcus faecalis";
RL Science 299:2071-2074 (2003).
DR ENBL; A0816955; AAO82308.1; -.
DR HSSP; P08837; 1GGR.
DR TIGR; EF2598; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008982; F:protein-N(P1)-phosphohistidine-sugar phospho. .; IEA.
DR GO; GO:0003351; F:sugar porter activity; IEA.
DR GO; GO:0009401; P:phosphoenolpyruvate-dependent sugar phospho. .; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR011055; Dup hybrid motif.
DR InterPro; IPR001996; P:trans_EIIB.
DR InterPro; IPR003352; P:trans_EIIC.
DR InterPro; IPR001127; PTS_EIIA.
DR InterPro; IPR011297; PTS_II_ABC_beta.
DR Pfam; PF00358; PTS_EIIA_1; 1.
DR Pfam; PF00367; PTS_EIIB; 1.
DR Pfam; PF02378; PTS_EIIC; 1.
DR ProDom; PD001476; P:trans_EIIB; 1.
DR ProDom; PD002243; PTS_EIIA; 1.
DR TIGRFAMs; TIGR00830; PTEA; 1.
DR TIGRFAMs; TIGR01995; PTS-II-ABC-beta; 1.
DR PROSITE; PS00371; PTS_EIIA_1; UNKNOWN_1.
DR PROSITE; PS01035; PTS_EIIB_CYS; 1.
KW Complete proteome.
SQ SEQUENCE 626 AA; 66244 MW; 0E0BF2C4220570B CRC64;

Query Match 30.3%; Score 715; DB 2; Length 626;
Best Local Similarity 37.4%; Pred. No. 6.8e-38;
Matches 166; Conservative 72; Mismatches 172; Indels 34; Gaps 7

QY 1 MAWVPSLVNGYDVAAATWAAGEMPWSLFGLDVNAQAQYQGTVLPVLVSVWILATIEKFLH 60
Db 181 MALVHPSI-----TALAGKTIISFAGLPVIIGSGYTSGLPFIILAVFAQSVYVERFK 232

QY 61 KRLKGTADFLIPLVLTLLLTGFLFIAGIPAMRWGVLAHSGQLYDRCGPGVGLLFG 120
Db 233 KVIPSEFLQICVPLAVFLIMAPVTFILAGITGVIGDWLGGYNALYAFSPITAGLNGS 292

QY 121 VYSPIVITGLHQSPPIIELEFLNFGSGSFIATASMANIAQAACLAFFVLAKSEKLKGLA 180
Db 293 LMQVLWMEGHWGFPVIMMLNLTOGGDTWPMLLPAVIAAGGAALAVFLLTKNVKLGUA 352

QY 181 GASGVSAVLGTEPAIFQVNLRWKPPFFIGTGAAGCALIALFNILKAVALGAA----- 234
Db 353 LSSSITITFITEPTVYGVTLPLKPTIAACIGGGIGGAFVAMNHVKNFTFGLVSLSLP 412

QY 235 GFLGVVSIDAPDMWMLVCAVTFIFAFGAAIAYGLVLRNGSIDPDATAPVPAQTGK 294
Db 413 GPFAETKDTAPMTGATGAGIAFIIFAVLT-----FVLRFEDQPNPE-----TATEK 460

FT DOMAIN 1 ? EIIB.
FT DOMAIN ? 480 EIIC.
FT MOD RES 26 26 Phosphocysteine (By similarity).
FT MOD RES 325 325 Phosphohistidine (By similarity).
SQ SEQUENCE 480 AA; 51326 MW; AB4E1D9785D84E47 CRC64;
Query Match 29.78; Score 701; DB 1; Length 480;
Best local similarity 51.4%; Pred. No. 4.1e-37; Indels 8; Gaps 4;
Matches 146; Conservative 49; Mismatches 81;
QY 1 MAMVPSLVNGYDVAATWAG-EMPMWSLFLGDLVAQAGYQGTLPVLVWVSWILATIEKFL 59
Db 194 MILVHPELMSAYDPKALEAGKEIPHNNLFLGLEINQVQYQGVLPMLVAVYILATIEKGL 253
QY 60 HKRUGTADFLITPVLLMLTGTFTTFAIGPAMRWGVDLAHGLQGLYDFGPGVGGILFG 119
Db 254 RKVIPTVLDNLLTPLLAILSTGFTTFSFVGLPTRLTGLYMSDGLTWLYEFGGATGGLIFG 313
QY 120 LVSPVITGLHQSFPPIELF---NOGGSFIFATASMANIAQGAACLAFFFLAK-SE 174
Db 314 LLYAPIVTGMHSHFIALETQIADSSSTGSGFIPPIATWSNIAQGAALAAAFIIEKNK 373
QY 175 KLKLAGASGVSALVIGITEPAIFGNLRLRPFFIGITAAIGGALLIALFNKKAVALGAA 234
Db 374 KLKGVASAGVSALLIGITEPAMFGVNLKLRYPFFIGALVSGIGSAYIAFFKVKALGTA 433
QY 235 GFLGVWSIDAPD--WMMLVCAVTFPTFAAGAAAYGLYLVRN 276
Db 434 GIPGFISISQNGWLMHYGIAMITAFIVAFGVTVALSRYKRYRN 477
RESULT 15
Q9KG19 PRELIMINARY; PRT; 630 AA.
AC Q9KG19;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PTS system, beta-glucoside-specific enzyme II, ABC component.
GN Name=BH0296;
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
Horikoshi K.;
RT "complete genome sequence of the alkaliphilic bacterium Bacillus
RL Nucleic Acids Res. 28:4317-4331 (2000).
DR EMBL; AP001508; BAB04015.1;
DR PIR; H83686; H83686.
DR HSSP; P20166; 1AX3.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008982; F:protein-N(Pi)-phosphohistidine-sugar phospho. .; IEA.
DR GO; GO:0005351; F:sugar porter activity; IEA.
DR GO; GO:0009401; P:phosphoenolpyruvate-dependent sugar phospho. .; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR011055; Dup hybrid motif.
DR InterPro; IPR001996; Ptrans_EIIB.
DR InterPro; IPR003352; Ptrans_EIIC.
DR InterPro; IPR001127; PTS_EIIA.
DR InterPro; IPR011297; PTS_II_ABC_beta.
DR Pfam; PF00358; PTS_EIIA_1; 1.
DR Pfam; PF00367; PTS_EIIB_1.
DR Pfam; PF02378; PTS_EIIC; 1.
DR ProDom; PD001476; Ptrans_EIIB; 1.
DR ProDom; PD00243; PTS_EIIA; 1.
DR TIGRFAMs; TIGR00830; PTBA; 1.
DR TIGRFAMs; TIGR01995; PTS-II-ABC-beta; 1.

DR PROSITE; PS00371; PTS_EIIA_1; 1.
DR PROSITE; PS01035; PTS_EIIB_CYS; 1.
SQ SEQUENCE 630 AA; 66978 MW; 6F0218011696ADD5 CRC64;
Query Match 29.4%; Score 695.5; DB 2; Length 630;
Best local similarity 32.8%; Pred. No. 1.2e-36;
Matches 152; Conservative 108; Mismatches 182; Indels 21; Gaps 5;
QY 2 AMVPPSLVNGYDVAATWAGEMPWSLFLGDLVAQAGYQGTLPVLVWVSWILATIEKFLHK 61
Db 185 ALIYPSIVELHDSALDV-----TFFGIPVLMNYTSTVFPIILLAVFAMSYVEKFCNK 236
QY 62 RLKGTADFLITPVLLMLTGTFTTFAIGPAMRWGVDLAHGLQGLYDFGPGVGGILFLV 121
Db 237 KIHEAVRNFVTPLLVLVIVPVTLIIILGPTGVILNGIASVIQEIFTFSPVIAAGI VAGI 296
QY 122 YSPVITGLHQSFPPIEL-ELFNOGGSFIFATASMANIAQGAACLAFFFLAKSEKLGLA 180
Db 297 WQVLVIFGIHWGIPIIILNNLSVRGEDIKAVAPAVFSQAGAALGVMLRTKNNKKALAA 356
QY 181 GASGVSAVLGITEPAIFGNLRLRPFFIGITAAIGGALLIALFNKKAVALGAAAGFLGVV 240
Db 357 GSTSITALEGITEPAVVGVTLPKKPFTMAVISAAGVAIVGHYGSVAVAFAPGLLTIP 416
QY 241 SI---DAPDMVFLVCAVTFPTFAAGAAAYGLYLVRNGSIDPDATAAPVAGTTKAEA 297
Db 417 IFYPEDGRGFVAFVIAIIISFVLA-----AVLTIVYGVKDPVDEDTLSNSSGSENEVR 471
QY 298 EAPAEFNSDSTIIQAPLTGREALALSSVSDAMFASGKLGSGVAIVPTKGQLVSPVSGKIIV 357
Db 472 BDDKKRPSASEIKSPKLGVEVPLTEVDHVFSSGAMGKGVAVRPKEGELVAPINGTVTS 531
QY 358 APPSGHAPAVRTKAEDGSNDILMHIGEDTVNLNGTHFNPLKKGDEVKAGELLCEFDID 417
Db 532 LFEYKHAIGITS---DNGTEIFIHVGIDTVQLKGEHFTSFTEQGEVAAGVLLLEFDVE 587
QY 418 AIKAGYEVVTPPIVVSNNYKTKGTPNTYGLGLEIAGANLLNVAK 460
Db 588 RITAAGYDVITPVLITNAKQFSNVQTTDKREVTSDDLIIHVIK 630
Search completed: October 29, 2004, 23:14:21
Job time : 201 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 29, 2004, 23:05:49 ; Search time 40 Seconds
(without alignments)
1125.736 Million cell updates/sec

Title: US-09-604-231-2
Perfect score: 2363
Sequence: 1 MAWVPSLVNGYDVAAATMAA.....IEAGANLLNVAKEAVPATP 468
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1058.5	44.8	651	S44257	phosphotransferase
2	953.5	40.4	664	B32243	phosphotransferase
3	759.5	32.1	627	H95951	fusion, PTS system
4	701	29.7	480	S33978	scra protein - Sta
5	695.5	29.4	630	H83686	PTS system, beta-g
6	695	29.4	617	AB1167	phosphotransferase
7	674	28.5	633	AC1436	PTS system, beta-g
8	673.5	28.5	632	C83724	PTS system, beta-g
9	673	28.5	632	S68599	phosphotransferase
10	673	28.5	655	C95220	trehalose PTS syst
11	667	28.2	634	AD1078	PTS system, beta-g
12	664	28.1	627	F95200	PTS system IABC C
13	660	27.9	627	F98067	phosphotransferase
14	659	27.9	479	F82432	PTS system, sucros
15	659	27.9	480	D90038	PTS system, sucros
16	658	27.8	705	A95084	phosphotransferase
17	657	27.8	609	I40406	beta-glucoside per
18	646	27.3	479	JQ0781	sucrose uptake pro
19	645	27.3	609	T47097	hypothetical prote
20	640	27.1	631	B42603	beta-glucoside-spe
21	628	26.6	617	AC1421	beta-glucoside-spe
22	612	25.9	628	D97073	PTS system, beta-g
23	588.5	24.9	636	D86807	hypothetical prote
24	580	24.5	618	AC1204	phosphotransferase
25	578.5	24.5	612	A97935	hypothetical prote
26	570.5	24.1	612	B95067	hypothetical prote
27	570	24.1	470	C69725	phosphotransferase
28	569.5	24.1	470	H33926	PTS system, trehal
29	567.5	24.0	625	C25977	phosphotransferase

30	509	21.5	640	2	AB1423	beta-glucoside-spe
31	496	21.0	475	2	C89813	hypothetical prote
32	489	20.7	456	2	S62331	phosphotransferase
33	487	20.6	458	2	H83881	PTS system, sucros
34	473	20.0	455	1	WQEBST	phosphotransferase
35	467	19.8	372	2	I39868	sac operon regulat
36	408	17.3	494	2	AG1231	PTS system trehalo
37	406	17.2	459	2	JU0293	levansucrase synth
38	396.5	16.8	494	2	AF1585	PTS system trehalo
39	390	16.5	460	2	A39938	phosphotransferase
40	367	15.5	473	2	A98281	trehalose specific
41	367	15.5	473	2	C65236	phosphotransferase
42	367	15.5	483	2	AI0449	protein-Npi-phosph
43	360	15.2	473	2	A86122	PTS system enzyme
44	348.5	14.7	681	2	A89781	hypothetical prote
45	345	14.6	665	2	B96970	PTS enzyme II, ABC

ALIGNMENTS

RESULT 1

S44257
phosphotransferase system enzyme II (EC 2.7.1.69), sucrose specific - Pediococcus pentose
C;Species: Pediococcus pentosaceus
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C;Accession: S44257
R;Leenhouts, K.K.J.; Bolhuis, A.A.; Kok, J.J.; Venema, G.G.
submitted to the EMBL Data Library, April 1994
A;Description: The sucrose and raffinose operons of Pediococcus pentosaceus PPE1.0.
A;Reference number: S44252
A;Accession: S44257
A;Molecule type: DNA
A;Residues: 1-651 <LEE>
A;Cross-references: UNIPROT:P43470; EMBL:Z32771; NID:G493728; PIDN:CAA83668.1; PID:G47596
C;Gene: scra
C;Superfamily: phosphotransferase system enzyme II sucrose-specific; phosphotransferase
C;Keywords: phosphotransferase
F;488-640/Domain: phosphotransferase system glucose-specific enzyme II, factor III homolo

Query Match	44.8%;	Score	1058.5;	DB 2;	Length	651;			
Best Local Similarity	46.7%;	Pred. No.	1.3e-64;						
Matches	221;	Conservative	78;	Mismatches	151;	Indels	23;	Gaps	5;
QY	1	MAWVPSLVNGYDVAAATMAAGEMPWWSLFGLDVAAGYQGTVPVLVWSWILATIEKFLH	60						
Db	194	MIMVLFSLVNGYSVATTMAAGRWYMNVEGLHVAQAGYQGQVLPVLGVAFILATLEKPFH	253						
QY	61	KRLKGTADFLITPVLTLTLLTGELTFTIAGPMRWYGDVLAHGLQGLYDFGPGVGLLFGI	120						
Db	254	KHKGAFDFTTTPMFAIVITGELTFTIIVGPVLRITVSDALTNGLVGLYNTSGWIGMIGFGL	313						
QY	121	VYSPVITIGLHQSPPIELELF-----NQGGSFIFATASMANIAQGAACLVAFVFLAKSEKL	176						
Db	314	LYSAIVITIGLHQTPAIETQLLANVAKTGGSFIPFVSVANIGQAALVAFVFLAKSEKL	373						
QY	177	KLAGASGVSAVLIGITEPAIFGNLRLRWPFPIGIGTAAIGCALIALPNIKAVLGAAGF	236						
Db	374	KALTSSAGVSALLGITEPAIFGNLRLRWPFPIGIGTAAIGCALIALPNIKAVLGAAGF	433						
QY	237	LGWVSIDAPDMVFLVCVAVTFETAFGAATAYGLYLVNRNSIDPDATAAPVAGTTKAE	296						
Db	434	IGFTISAKSIPAFMLSAVISFVVAFTPTFIY-----AKRTLGDRDQVKSAPTSTV---	486						
QY	297	AEAPAEFSNDSTIIQAPLTGEAIALSSVSDAMFASGKLGSGVAIVPTKQGLVSPVSGKIV	356						
Db	487	-----INVNDEIISAPVTGASESLKQVNDQVPSAEIMKGAAIVPSSDQVVPADGVIT	540						
QY	357	VAFPSGHAFVTRKAEDGSNNVDILMHIGFDTVNLNGHTFNPLKQGDVEKAGELICEPDI	416						
Db	541	VTYDSHAYGIKTTA-----GAELIHLGLDVTNNGEHTTNVQKGVDTVHQDILLGTFTDI	596						

Science 294, 849-852, 2001
A;Authors: Krefit, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlutterm, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,
A;Title: Comparative genomics of *Listeria* species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AC1436
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-633 <GLA>
A;Cross-references: UNIPROT:Q92FS7; GB:AL592022; PIDN:CAC95259.1; PID:g16412447; GSPDB:G
A;Experimental source: strain Clp11262
C;Genetics:
C;Superfamily: phosphotransferase system enzyme II sucrose-specific; phosphotransferase
Query Match 28.5%; Score 674; DB 2; Length 633;
Best Local Similarity 36.2%; Pred. No. 2.1e-38;
Matches 163; Conservative 75; Mismatches 179; Indels 30; Gaps 10;
QY 2 AMVFPSSL-VNGYDVAATMAAG---EMPMSLF-GLDVAQAGYQGTVPVLVLSWILATI 55
DB 181 ALVYPTLAGISAGDPVYTFAGTIFESPVHTFGLIPVILMSYASSVPIIATYFGSKV 240
QY 56 EKFLHKLKGTADFLITPVLLTLLTGTFTFIAIGPAMRWGDVLAHGIQGLYDFGGPYGG 115
DB 241 EKGFKKIIPDVVKTFFVFCVLLVVPVTFVIGPIATWAGQLLCAGTIWYNLSPVAVG 300
QY 116 LFLGLVSPVITGLHQSPPIEL-ELFNQGGSFIFATASMANIAQGAACLAIVFLAKSE 174
DB 301 LILGFGVQVFFVFLGLHWGLIPVAINNLTGSDPVLAMFGASFAQIAGVLAIVFFKTRNK 360
QY 175 KLKGLAGASGVAVLIGTEPAIFGNRLRWPFFIGITGTAIGGALIALENIKAVAGAA 234
DB 361 KIKLSIPAFISGIFGVTEPAIYVTLPLKPFINSIAGAVGGIIGFVSAKVIMGL 420
QY 235 GFLGVVSIDAP-----DMVPLVCVAVTFFTAAGAAIAGLYLVRRNGSIDPDATAAPVP 289
DB 421 GFLGPNFFQPGAGITSAPWVVIIVISFI-----LGFILTYVAGFKDP--AEAVV 470
QY 290 AGTTKAAEAPAEFNSDSTIIQAPITGEAIALSSVDSAMFASGKLGSGVAIVPTGQLVS 349
DB 471 EETNTEGETLIE----RETIPAPVVGRIIVTLADVKDEAFSSGALGKGVAILPTVGRVVA 526
QY 350 PVSGKIVVAFPSGHAPAVTKAEDGNSVDILMHIGDFTVNLGTHFNPLKKGDSVKAGE 409
DB 527 PAAGTVTTFPTGHAIGITT--NDGA--EVLHIGMDTVQLEGKFTTAHVKGQDVIEKQ 582
QY 410 LLCEFDIDAIKAAGYEVTTPIVVSN 434
DB 583 LUTEFDIEGKAAGYDVTTTPVVVTN 607
RESULT 8
C83724
PRS system, beta-glucoside-specific enzyme II, ABC component bglP [imported] - Bacillus
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C;Accession: C83724
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: AB3650; MUID:20512582; PMID:11058132
A;Accession: C83724
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-636 <STO>
A;Cross-references: UNIPROT:Q9KF90; GB:AP001509; GB:BA000004; NID:g10173176; PIDN:BA043
A;Experimental source: strain C-125
C;Genetics:
C;Superfamily: phosphotransferase system enzyme II sucrose-specific; phosphotransferase
Query Match 28.5%; Score 673.5; DB 2; Length 636;

Best Local Similarity 34.0%; Pred. No. 2.2e-38;
Matches 163; Conservative 78; Mismatches 193; Indels 45; Gaps 9;
QY 2 AMVFPSSLVNGYDVAATMAAGEMPMSLF-----GLDVAQAGYQGTVPVLV 47
DB 183 SLVYPTLV-----VLTEGE-PLYTLFTGTIFESPVHTFGLIPVILMSYATSVIPIIL 234
QY 48 VSWILATIEKFLHKLKGTADFLITPVLLTLLTGTFTFIAIGPAMRWGDVLAHGLGGLY 107
DB 235 AAYFASKVEARLARKIIPDVVKTFELVPFFTLVIVPLTFFIVIGPIATWAGQLLGQFTLWVY 294
QY 108 DFGPVGGLLFLVYSPVITGLHQSPPIEL-ELFNQGGSFIFATASMANIAQGAACLA 166
DB 295 NLSPIIAGAFGLGFWQVFFVFLGHGWLIPAINNLVVGSDPVLAMVFAASFAQIGAVAA 354
QY 167 VFPLAKSEKLGLAGASGSAVLGITEPAIFGNRLRWPFFIGITGTAIGGALIALENI 226
DB 355 VWLKIKQKQKVTLSVPAFISGIFGVTEPAIYGVTLPLKRPFIISCIAAVAGGAILGLFRS 414
QY 227 KAVAGAGFLGVVSI-----DAPDMVMP-LVCVAVTFFIAGAAIAYGLIVRRNGSIDP 281
DB 415 QGYIIGLGIIFGIPSLHPADGMDAGFWGIVIAVVAVFV-LGFLTITLFLGLKSGNASDEQ 473
QY 282 DATAAPVAGTTKAAEAPAEFNSDSTIIQAPITGEAIALSSVDSAMFASGKLGSGVAIV 341
DB 474 TETKAHTSTGTGEKEE-----ISSPFNGSVITLSEIKDEAFSSGALGEGIAIE 521
QY 342 PTKGQLVSPVSGKIVVAPPSPGHAFAPVTKAEDGNSVDILMHIGDFTVNLGTHFNPLKKQ 401
DB 522 PSEKGLSPVSGMWTALYPTTHALGIIT---DRGAELLIHIGLDTVQLDGKFFTAHTIQ 577
QY 402 GDEVKAGELLCEFDIDAIKAAGYEVTTPIVVSNNYKKTGPVNTYGLGTEAGANLNAVAK 460
DB 578 GAQVEKGLLITFEFDIKEIKAAGYAVTTPVIVTNHKKYQGLFLTDKQQVNAGRDLLELTR 636
RESULT 9
S68599
phosphotransferase system enzyme II (EC 2.7.1.69), sucrose-specific - Streptococcus sobri
N;Alternate names: sucrose-specific enzyme II
C;Species: Streptococcus sobrinus
A;Variety: strain 6715
C;Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 07-May-1999
C;Accession: S68599
R;Chen, Y.Y.M.; Lee, L.N.; LeBlanc, D.J.
Infect. Immun. 61, 2602-2610, 1993
A;Title: Sequence analysis of scrA and scrB from Streptococcus sobrinus 6715.
A;Reference number: S68598; MUID:93273516; PMID:8500898
A;Accession: S68599
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-632 <CHE>
A;Cross-references: EMBL:L06791
C;Genetics:
A;Gene: scrA
C;Superfamily: phosphotransferase system enzyme II sucrose-specific; phosphotransferase
C;Keywords: phosphotransferase; sugar transport system
F;480-632/Domain: phosphotransferase system glucose-specific enzyme II, factor III homol
Query Match 28.5%; Score 673; DB 2; Length 632;
Best Local Similarity 36.4%; Pred. No. 2.4e-38;
Matches 168; Conservative 76; Mismatches 197; Indels 20; Gaps 8;
QY 1 MAMVFPSSLVNGYDVAATMAAGEMPMSLFGLDVAQAGYQGTVPVLVIVSWILATIEKFLH 60
DB 191 LMAVNPALPNAYAV---ASGDAKALTFEGF-IPVVGYQGTVPFAFFVGMIGARLENWLH 245
QY 61 KRLKGTADFLITPVLLTLLTGTFTFIAIGPAMRWGDVLAHGLQGLYDFGPGVGGLLFGL 120
DB 246 KRVPEALDLITPFLTFLVMSILGLFAIGGVFHSVETVLAATENWIALPFGIAGIILGG 305
QY 121 VYSPVITGLHQSPPIELFLFNQGGSFIF-ATASMANIAQGAACLAIVFFLAKSEKLGKL 179

Db 306 LQQVIVVTGVHHIFNFLETQQLAETKANPFPNLLSAATAGQGVAVLAVAVKTKSAKLKAL 365
QY 180 AGAGSVSAVIGITEPAIFGVNLRWPFPIGCTAAIGGALIALFNKAKAVALGAAGFLGV 239
Db 366 AYPSSALSAALGITEPAIFGVNLRWPFPIGCTAAIGGALIALFNKAKAVALGAAGFLGV 425
QY 240 VSDADPMVMFLYCAVUTPFIAGCAIAYGLYLVRNGSIDDPDATAAPVPAGTTKAEAEA 299
Db 426 LFLNSQMPYIVSIITVACAIAP--ALTY--YFGYADKEEDVSARKPEAPAAAPVAETET 481
QY 300 PARFNSDSTIIQAPLTGEALSSVSDAMFASGKLGSGVAIVPTKQGLVSPVSGKIIVAF 359
Db 482 KSE-----VIASFLQGEAVELSKVNDPVFSSAMGKIAVSPGNTVYSPVNGTVQIAF 535
QY 360 PSCHAFVATKAEAGNSVDILMHIGFDTVNLNGTHFNPLKQKQDEVKAGBELLCEFFIDAI 419
Db 536 ETCHAYGL--KSDNGA--EVLHVGIDTVSMNGTGFQDKVAANQTVKGVGVLTGTFSAKI 591
QY 420 KAAGYEVTPIVSNYKKTGPVNTYGLGEIEAGANLLINAK 460
Db 592 ABAGLDDTTWVIITNTADYSEVKPLAAGQLAHAGLLELNK 632
RESULT 10
C95220
trehalose PTS system, IIABC components [imported] - Streptococcus pneumoniae (strain TIC
C:Species: Streptococcus pneumoniae
C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
C:Accession: C95220
R:Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Bisen, J.A.; Read, T.D.; Peterson, S.; Heid
on, J.D.; Unayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapfel,
nson, T.; Hickey, E.K.; Holt, I.B.
Science 293, 498-506, 2001
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A:Reference number: A95000; MUID:21357209; PMID:11463916
A:Accession: C95220
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-655 <KUR>
A:Cross-references: UNIPROT:Q97NW9; GB:AE005672; PIDN:AAK75956.1; PID:g14973388; GSPDB:C
A:Experimental source: strain TIGR4
C:Genetics:
A:Gene: SP1884
C:Superfamily: phosphotransferase system enzyme II sucrose-specific; phosphotransferase
Query Match 28.5%; Score 673; DB 2; Length 655;
Best Local Similarity 35.3%; Pred. No. 2.5e-38;
Matches 170; Conservative 87; Mismatches 173; Indels 52; Gaps 16;
QY 1 MAMVPPSLVNGYDVAAATMAAGEMP--MMSLFGLDVAQAGYQGTLPVLVWVSMILATIEKF 58
Db 195 ICLVSPQLLNAYAVASTPAADIAANWNNFGYFTVNRIGYQAVIPALLAGLSLSYLIF 254
QY 59 LHKRLKGTADFLITPVLILLTGTFTTAIGPAMRWGVDVLAHQGLYDFG--GPVG-- 114
Db 255 WHKHIPVISMIFVFLSLIFALIAHTVLGP-----IGWTIGQGLSSVVLVAGLTGPKWL 310
QY 115 -GLLFLGVSPVITGLHOSPPPIELF--NOGGSFIFATASMANIAQGAACLAFFLAK 172
Db 311 FGAIFGALYAPFVITGLHMTNADTQIADAGGTALWPMIALSNIAQGSVAFFAYFMR 370
QY 173 -SEKLKLAGASGSVAVLGITEPAIFGVNLRWPFPIGCTAAIGGALIALFNKAKAVAL 231
Db 371 HDREAQVSLPATISAYLGVTEPALFGVNVKVIYFFVAGMTGSALAGMLSTFVNTAASI 430
QY 232 GAAGFLGVSDADPMVMF-----LVCAVUTPFIAGCAIAYGLYLVRNGSIDDPDATAAP 287
Db 431 GIGGLPGILSIQPTMLFPFAGTMLVAVVPMLLTF-----FFRKAGLFTK----- 475
QY 288 VPAGTTKAEAE-----APAEFNSDSTI-----IQAPLTGEALSSVSDAMFASGKLGSG 337
Db 476 -TEGDTNLQAEFVQAEAEFVNHPEVELTSVELISPTGQVKELSQATDPIFASGVMGQG 534

QY 338 VAIVPTKQGLVSPVSGKIIVVAPPSCHAFVATKAEAGNSVDILMHIGFDTVNLNGTHFNP 397
Db 535 LVTEPSQGLTEPVGNTVTLVFFTKAIGT--VSDEG--VELLIHIGMDTVLGDGKGFES 590
QY 398 LKQKQDEVKAGBELLCEFFIDAIKAAGYEVTPFIWVSN--YKKTGPVNTYGLGIEIAGAN 454
Db 591 LVVQGHVTVGQQLIRFDMVIVKAGLVTEPFIINQDAYTATIP--GTPT--TIQAGAS 648
QY 455 LL 456
Db 649 LM 650
RESULT 11
AD1078
PTS system, beta-glucosides specific enzyme IIABC homolog lmo0027 [imported] - Listeria
C:Species: Listeria monocytogenes
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: AD1078
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker,
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Psihi, H.;
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Krefte, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, B.; Maitournam, A.; Mat
ok, C.; Schluter, T.; Simoes, N.; Trierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AD1078
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-634 <GLA>
A:Cross-references: UNIPROT:Q8YAT6; GB:NC_003210; PIDN:CAC98242.1; PID:g16409386; GSPDB:C
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: lmo0027
C:Superfamily: phosphotransferase system enzyme II sucrose-specific; phosphotransferase
Query Match 28.2%; Score 667; DB 2; Length 634;
Best Local Similarity 34.6%; Pred. No. 6.2e-38;
Matches 156; Conservative 81; Mismatches 172; Indels 42; Gaps 10;
QY 2 AMVFPSSL--VNGYDVAAATMAAG--EMPWMSLF--GLDVAQAGYQGTLPVLVWVSMILATI 55
Db 182 ALVYFTLAGITAGDPHYLTFAGTIFESPIHVTFLGIFVILMSYASSVPIIVATYFGSKV 241
QY 56 EKFLHKLKGTADFLITPVLILLTGTFTTAIGPAMRWGVDVLAHQGLYDFGPGVG 115
Db 242 EKGKKIIPDVIKTPVPFCTLLIVVPITFIVGPIATWAGQLLGGATVWVNLSPING 301
QY 116 LLFGLVYSPVITGLHOSPPPIEL--ELFNQGSFIFATASMANIAQGAACLAFFLAKSE 174
Db 302 LILGFMQVQVIFGLHMLGVPAVINLTVLGHDPILAMTFGASFAQIGAVLAVFFKSRNK 361
QY 175 KLKLAGASGSVAVLGITEPAIFGVNLRWPFPIGCTAAIGGALIALFNKAKAVALGAA 234
Db 362 KIKLSIPAFISFGVTEPAIYGVTPKPKPFINSCIAAGGIGGGIIGFAGSQTYIMGSL 421
QY 235 GFLGVVSDADP-----DMVMFLVCAVUTPFIAGCAIAYGL-----YLVNRNGSIDPDA 283
Db 422 GIFGLNPFKPGSGISGFEFWVVIATVIFLITVVGFKDPADVVPVQSNTIVE--- 478
QY 284 TAAPVAGTTKAEAEAPAFNSDSTIIQAPLTGEALSSVSDAMFASGKLGSGVAIVPT 343
Db 479 -----GETLIERET-----IPAPVVGIVTLADVKDEAFSSGALGKGVAIPT 521
QY 344 KGQLVSPVSGKIIVVAPPSCHAFVATKAEAGNSVDILMHIGFDTVNLNGTHFNPLKQGD 403
Db 522 VGRVAPAGTVTTFPTGHAIGTITK--DGA--EVLIIHIGMDTVQLEKGFTHFVKQGD 577
QY 404 EVKAGBELLCEFFIDAIKAAGYEVTPFIWVSN 434
Db 578 VIEKQLLTFFEDIGIKAGYDVTPVVTN 608

A;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A;Reference number: A97872; PMID:21429245; PMID:11544234

A;Accession: E98067
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-627 <KUR>
A;Cross-references: UNIPROT:Q8DMS8; GB:AE007317; PIDN:AAL00370.1; PID:gl5459232; GSPDB:G
C;Genetics:
C;Superfamily: phosphotransferase system enzyme II sucrose-specific; phosphotransferase
C;Keywords: phosphotransferase

Query Match 27.9%; Score 660; DB 2; Length 627;
Best Local Similarity 34.3%; Pred. No. 1.8e-37;
Matches 159; Conservative 92; Mismatches 187; Indels 26; Gaps 10;

QY 1 MAMVPSLVNGYDVAAATMAAGEMPMSLFLGLDVAQAGYQGTVPVLVVSMTLATEIKFLH 60
DB 182 MMLVSGSLPNAWA--OGEVMTAMNFGP-IPVVGLOGSVLPFAFIIGVVGAKFEKAVR 237
QY 61 KRLKGTADFLITPVLTLTLLTGFTLTFIAIGPAMRWGVDVLAHGLQGLYDFGPGVGLLFLGL 120
DB 238 KVPDVIDLLVTPFVTLVMSILGLFVIGPVFHVVENVILIATKAILSMPPFLGGLIGG 297
QY 121 VYSPVITGLHQSPPIELFENQGGSPF-ATASMANIAGAACLAFFFLAKSEKLKGL 179
DB 298 VHQLIVSGVHHIFNLLEVLAAADHANFPNFIITAAATAQGAATVAVGVTKNPKLKITL 357
QY 180 AGAGSVNAVIGITPAIFGVNLRWRPFPGIGTAAIGGALIALFNKAKALGAAGFLG- 238
DB 358 APPAALSFLGITPAIFGVNLRWRPFPGIGTAAIGGALIALFNKAKALGAAGFLG- 417
QY 239 VVSDADPMWFLVCAVVTFFIAFGAATAYGLYLVRRNGSIDPDATAAPVPAGTTKAAE 298
DB 418 MLYVNGQLPQYLLVAVSFALGTYMFGY-----EDEVATAAKAEVAEKEE 470
QY 299 -APAEFNSDSTIIQAPLTGEATLSSVDAMFASGLSGVAIVPTKQLVSPVSGKIIV 357
DB 471 VAPAAALQNETLV--TPVIGDVVALADVNDPVFSSGAMGGIIVKPSQGVVYAPADAESVI 528
QY 358 APPSGHAFVTRKAEDEGNSVDILMHIGFTVNLNTHFNPLKKQGVKAGELLCEFFDID 417
DB 529 APPTGHAFGLKTR---NGAEVLHVGIDTVSMNGDGFEAKVQGNKVKAGDVLGTFDSN 584
QY 418 AIKAGYEVVTPPIVWSN---YKKTGPVNTYGLGEIEAGANLLNV 458
DB 585 KIAAAGLDDTTWIVTNTGDYASVAPVAT---GSVSKGDVAVIEV 625

RESULT 14

F82432
PTS system, sucrose-specific IIBC component VCA0653 [imported] - Vibrio cholerae (strain
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
R;Heidelberg, J.F.; Eissen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000

A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;Reference number: A82035; PMID:20406833; PMID:10952301
A;Accession: F82432
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-479 <HEI>
A;Cross-references: UNIPROT:Q9KLT8; GB:AE004395; GB:AE003893; NID:99658068; PIDN:AF9655

C;Genetics:
A;Map position: 2
C;Superfamily: phosphotransferase system sucrose-specific enzyme II, factor II
Query Match 27.9%; Score 659; DB 2; Length 479;

RESULT 12

F95200
PTS system IIBC components [imported] - Streptococcus pneumoniae (strain TIGR4)
C;Species: Streptococcus pneumoniae
C;Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
R;Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eissen, J.A.; Read, T.D.; Peterson, S.; Heid
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapfle,
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001

A;Authors: Iofutus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A;Reference number: A95000; PMID:21357209; PMID:11463916
A;Accession: F95200
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-627 <KUR>
A;Cross-references: UNIPROT:Q9TPB8; GB:AE005672; PIDN:AAK7599.1; PID:gl4973217; GSPDB:G
A;Experimental source: strain TIGR4
C;Genetics:
C;Superfamily: phosphotransferase system enzyme II sucrose-specific; phosphotransferase

Query Match 28.1%; Score 664; DB 2; Length 627;
Best Local Similarity 34.3%; Pred. No. 9.8e-38;
Matches 159; Conservative 92; Mismatches 187; Indels 26; Gaps 10;

QY 1 MAMVPSLVNGYDVAAATMAAGEMPMSLFLGLDVAQAGYQGTVPVLVVSMTLATEIKFLH 60
DB 182 MMLVSGSLPNAWA--OGEVMTAMNFGP-IPVVGLOGSVLPFAFIIGVVGAKFEKAVR 237
QY 61 KRLKGTADFLITPVLTLTLLTGFTLTFIAIGPAMRWGVDVLAHGLQGLYDFGPGVGLLFLGL 120
DB 238 KVPDVIDLLVTPFVTLVMSILGLFVIGPVFHVVENVILIATKAILSMPPFLGGLIGG 297
QY 121 VYSPVITGLHQSPPIELFENQGGSPF-ATASMANIAGAACLAFFFLAKSEKLKGL 179
DB 298 VHQLIVSGVHHIFNLLEVLAAADHANFPNFIITAAATAQGAATVAVGVTKNPKLKITL 357
QY 180 AGAGSVNAVIGITPAIFGVNLRWRPFPGIGTAAIGGALIALFNKAKALGAAGFLG- 238
DB 358 APPAALSFLGITPAIFGVNLRWRPFPGIGTAAIGGALIALFNKAKALGAAGFLG- 417
QY 239 VVSDADPMWFLVCAVVTFFIAFGAATAYGLYLVRRNGSIDPDATAAPVPAGTTKAAE 298
DB 418 MLYVNGQLPQYLLVAVSFALGTYMFGY-----EDEVATAAKAEVAEKEE 470
QY 299 -APAEFNSDSTIIQAPLTGEATLSSVDAMFASGLSGVAIVPTKQLVSPVSGKIIV 357
DB 471 VAPAAALQNETLV--TPVIGDVVALADVNDPVFSSGAMGGIIVKPSQGVVYAPADAESVI 528
QY 358 APPSGHAFVTRKAEDEGNSVDILMHIGFTVNLNTHFNPLKKQGVKAGELLCEFFDID 417
DB 529 APPTGHAFGLKTR---NGAEVLHVGIDTVSMNGDGFEAKVQGNKVKAGDVLGTFDSN 584
QY 418 AIKAGYEVVTPPIVWSN---YKKTGPVNTYGLGEIEAGANLLNV 458
DB 585 KIAAAGLDDTTWIVTNTGDYASVAPVAT---GSVAKGDVAVIEV 625

RESULT 13

E98067
phosphotransferase system enzyme II (EC 2.7.1.69) scra [imported] - Streptococcus pneum
C;Species: Streptococcus pneumoniae
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
R;Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; B
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
Y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A;Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 29, 2004, 23:14:30 ; Search time 130 seconds
(without alignments)
1167.177 Million cell updates/sec

Title: US-09-604-231-2

Perfect score: 2363

Sequence: 1 MAMVFPISLVNGYDVAAATWAA.....IEAGANLLNVAKEAVPATP 468

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1370721 seqs, 324215800 residues

Total number of hits satisfying chosen parameters: 1370721

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

```
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09D_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US10F_PUBCOMB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US10G_PUBCOMB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US10H_PUBCOMB.pep.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2363	100.0	468	15	US-10-450-055-10
2	2363	100.0	661	9	US-09-738-626-6404
3	695	29.4	617	15	US-10-282-122A-60875
4	673	28.5	655	16	US-10-474-776-375
5	660	27.9	627	9	US-09-815-242-13467
6	585	24.8	381	16	US-10-474-776-367
7	578.5	24.5	612	15	US-10-282-122A-73843
8	570.5	24.1	612	16	US-10-474-776-253
9	564.5	23.9	620	15	US-10-282-122A-74439
10	492	20.8	379	15	US-10-282-122A-70209
11	468.5	19.8	683	9	US-09-738-626-6961
12	468.5	19.8	683	15	US-10-450-055-26
13	348.5	14.7	681	15	US-10-282-122A-43922

14	345.5	14.6	679	9	US-09-815-242-5658	Sequence 5658, Ap
15	345.5	14.6	681	9	US-09-815-242-12270	Sequence 12270, A
16	345	14.6	665	15	US-10-282-122A-51493	Sequence 51493, A
17	338.5	14.3	484	9	US-09-815-242-10809	Sequence 10809, A
18	338.5	14.3	484	15	US-10-282-122A-56917	Sequence 56917, A
19	336.5	14.2	474	15	US-10-282-122A-70588	Sequence 70588, A
20	331.5	14.0	453	15	US-10-282-122A-55853	Sequence 55853, A
21	329.5	13.9	526	15	US-10-282-122A-57960	Sequence 57960, A
22	323	13.7	451	15	US-10-282-122A-57628	Sequence 57628, A
23	323	13.7	484	9	US-09-815-242-12272	Sequence 12272, A
24	323	13.7	484	15	US-10-282-122A-44138	Sequence 44138, A
25	322	13.6	650	15	US-10-282-122A-75030	Sequence 75030, A
26	321.5	13.6	453	15	US-10-282-122A-75796	Sequence 75796, A
27	319.5	13.5	429	15	US-10-282-122A-73368	Sequence 73368, A
28	319.5	13.5	650	15	US-10-282-122A-75740	Sequence 75740, A
29	318.5	13.5	651	15	US-10-282-122A-59437	Sequence 59437, A
30	316.5	13.4	466	15	US-10-282-122A-72912	Sequence 72912, A
31	315.5	13.4	677	15	US-10-282-122A-53222	Sequence 53222, A
32	308.5	13.1	677	15	US-10-282-122A-70806	Sequence 70806, A
33	306.5	13.0	657	15	US-10-282-122A-71894	Sequence 71894, A
34	305.5	12.9	455	9	US-09-815-242-13794	Sequence 13794, A
35	304.5	12.9	454	15	US-10-282-122A-45789	Sequence 45789, A
36	300	12.7	687	9	US-09-815-242-5822	Sequence 5822, Ap
37	300	12.7	719	9	US-09-815-242-12942	Sequence 12942, A
38	299.5	12.7	450	15	US-10-282-122A-71596	Sequence 71596, A
39	299	12.7	676	15	US-10-282-122A-52557	Sequence 52557, A
40	299	12.7	726	16	US-10-474-776-280	Sequence 280, App
41	298.5	12.6	648	9	US-09-741-669-330	Sequence 330, App
42	298.5	12.6	648	9	US-09-815-242-10089	Sequence 10089, A
43	298.5	12.6	648	15	US-10-282-122A-42726	Sequence 42726, A
44	298	12.6	665	15	US-10-282-122A-57245	Sequence 57245, A
45	293.5	12.4	454	15	US-10-282-122A-45636	Sequence 45636, A

ALIGNMENTS

RESULT 1
US-10-450-055-10
; Sequence 10, Application US/10450055
; Publication No. US20040043953A1
; GENERAL INFORMATION:
; APPLICANT: BASF Aktiengesellschaft
; TITLE OF INVENTION: NO. US20040043953A1el genes of Corynebacterium
; FILE REFERENCE: 936 2000
; CURRENT APPLICATION NUMBER: US/10/450,055
; CURRENT FILING DATE: 2003-06-10
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Vers. 2.0
; SEQ ID NO 10
; LENGTH: 468
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-10-450-055-10

Query Match 100.0%; Score 2363; DB 15; Length 468;
Best Local Similarity 100.0%; Pred. No. 1e-195;
Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MAMVFPISLVNGYDVAAATWAAAGEMPWSLFLGDLVDAQAGYQGTVPVLVSVVSLATIEKFLH	60
Db	1	MAMVFPISLVNGYDVAAATWAAAGEMPWSLFLGDLVDAQAGYQGTVPVLVSVVSLATIEKFLH	60
Qy	61	KRLKGTADFLITPVLITLLTGFTFFIAGPMRWVGDVLAHGLQGLDFFGVPVGLLFLGL	120
Db	61	KRLKGTADFLITPVLITLLTGFTFFIAGPMRWVGDVLAHGLQGLDFFGVPVGLLFLGL	120
Qy	121	VYSPDIVTGLHOSPPPIELELFNQGGSFIATASMANIAQGAACIAVFFLAKSKLKGIA	180
Db	121	VYSPDIVTGLHOSPPPIELELFNQGGSFIATASMANIAQGAACIAVFFLAKSKLKGIA	180
Qy	181	GASGVSAVLGITEPAITGVNLRLEWPFIFIGTAAIGGALITALENIKAVLGAAGFGLV	240

```

Db 181 GASGSAVLGITPEALFGVNLRLRWPFFIGTAAIGGALLIALFNKAKVALGAAGFLGVV 240
QY 241 SIDAPDMWFLVCAVVTFFIAFGAAIAYGLYLVRNGSIDPDATAVPVPGATTKAEAEAP 300
Db 241 SIDAPDMWFLVCAVVTFFIAFGAAIAYGLYLVRNGSIDPDATAVPVPGATTKAEAEAP 300
QY 301 AEFSDNSTIIQAPLTGEAIALSSVSDAMFASGKLGSGVAIVPTKGQLVSPVSGKIVVAFP 360
Db 301 AEFSDNSTIIQAPLTGEAIALSSVSDAMFASGKLGSGVAIVPTKGQLVSPVSGKIVVAFP 360
QY 361 SGHAFVRTKAEKGSNDVILMHIGFDTVNLNGTHFNFLKQGGDEVKAGELLCEFDIDAIAK 420
Db 361 SGHAFVRTKAEKGSNDVILMHIGFDTVNLNGTHFNFLKQGGDEVKAGELLCEFDIDAIAK 420
QY 421 AAGYEVTTPVWSNYKKTGPVNTYGLGEIEAGANLLNVAKEAVPATP 468
Db 421 AAGYEVTTPVWSNYKKTGPVNTYGLGEIEAGANLLNVAKEAVPATP 468

RESULT 2
US-09-738-626-6404
; Sequence 6404, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR FILING DATE: 1999-12-16
; PRIOR FILING DATE: 1999-12-16
; PRIOR FILING DATE: 2000-04-07
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: Patent in ver. 3.0
; SEQ ID NO 6404
; LENGTH: 661
; TYPE: PPT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6404

Query Match 100.0%; Score 2363; DB 9; Length 661;
Best Local Similarity 100.0%; Pred. No. 1.7e-195;
Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAMVPSLVNGYDVAAATWAAGEMPWMSLFGLDVAQAGYQGTVPVLVWSWILATIEKFLH 60
Db 194 MAMVPSLVNGYDVAAATWAAGEMPWMSLFGLDVAQAGYQGTVPVLVWSWILATIEKFLH 253
QY 61 KKLKGTADFLITPVLTLLTGTFTTATGPMRWGDVLAHGLQGLYDFGPGVGLLFG 120
Db 254 KKLKGTADFLITPVLTLLTGTFTTATGPMRWGDVLAHGLQGLYDFGPGVGLLFG 313
QY 121 VYSPVITGLHQSFPPIELFNQGSFTFATASMANIAGAACLVFFLAKSEKLGIA 180
Db 314 VYSPVITGLHQSFPPIELFNQGSFTFATASMANIAGAACLVFFLAKSEKLGIA 373
QY 181 GASGSAVLGITPEALFGVNLRLRWPFFIGTAAIGGALLIALFNKAKVALGAAGFLGVV 240
Db 374 GASGSAVLGITPEALFGVNLRLRWPFFIGTAAIGGALLIALFNKAKVALGAAGFLGVV 433
QY 241 SIDAPDMWFLVCAVVTFFIAFGAAIAYGLYLVRNGSIDPDATAVPVPGATTKAEAEAP 300

```

```

Db 434 SIDAPDMWFLVCAVVTFFIAFGAAIAYGLYLVRNGSIDPDATAVPVPGATTKAEAEAP 493
QY 301 AEFSDNSTIIQAPLTGEAIALSSVSDAMFASGKLGSGVAIVPTKGQLVSPVSGKIVVAFP 360
Db 494 AEFSDNSTIIQAPLTGEAIALSSVSDAMFASGKLGSGVAIVPTKGQLVSPVSGKIVVAFP 553
QY 361 SGHAFVRTKAEKGSNDVILMHIGFDTVNLNGTHFNFLKQGGDEVKAGELLCEFDIDAIAK 420
Db 554 SGHAFVRTKAEKGSNDVILMHIGFDTVNLNGTHFNFLKQGGDEVKAGELLCEFDIDAIAK 613
QY 421 AAGYEVTTPVWSNYKKTGPVNTYGLGEIEAGANLLNVAKEAVPATP 468
Db 614 AAGYEVTTPVWSNYKKTGPVNTYGLGEIEAGANLLNVAKEAVPATP 661

RESULT 3
US-10-282-122A-60875
; Sequence 60875, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR FILING DATE: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-06,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR FILING DATE: 2000-05-26
; PRIOR FILING DATE: 2000-09-06
; PRIOR FILING DATE: 2000-09-06
; PRIOR FILING DATE: 2000-09-09
; PRIOR FILING DATE: 2000-09-09
; PRIOR FILING DATE: 2000-10-23
; PRIOR FILING DATE: 2000-11-27
; PRIOR FILING DATE: 2000-12-22
; PRIOR FILING DATE: 2001-02-09
; PRIOR FILING DATE: 2001-02-09
; PRIOR FILING DATE: 2001-02-16
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 60875
; LENGTH: 617
; TYPE: PPT
; ORGANISM: listeria monocytogenes
US-10-282-122A-60875

Query Match 29.4%; Score 695; DB 15; Length 617;
Best Local Similarity 34.8%; Pred. No. 3.2e-51;
Matches 154; Conservative 82; Mismatches 162; Indels 44; Gaps 6;

QY 2 AMVPSLVNGYDVAAATWAAGEMPWMSLFGLDVAQAGYQGTVPVLVWSWILATIEKFLH 61
Db 185 ALVYPTMNLNEGAIHFLQIP-----VLMYSYSPFVPIILAVNLSILERFLNS 236

```

```
QY 62 RUKGTADFLITPLVLLLTGFTTFAIGPAMRWGDLAHGLQGLYDFGPGVGLLFGIV 121
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 237 KIHEAAKTELTWICMLIVPLITFLAFLGFTFIQGLASGYTFYINLSPIVAGAFMGAP 296

QY 122 YSPVITGLHQSPFIEL-ELFNQGSFTFATASMANIAQGAACLAFFLAKSEKLGIA 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 297 WQVLVIFGHWGFVPTIMNLSRYGRDWTIAMVGPSTFAQAGASLGVFLTKKPKVAIA 356

QY 181 GASGVSALVIGITEPAIFGNLRWPFRTIGTAAIGGALIALENIKAVAGAGFL--- 237
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 357 GSAALTGGFTGTEPSYGLVKPKFPVJASTAGAIAGGAAGSSGAANAIPGILTL 416

QY 238 -----GVVSDIDAPDMVFLVCAVVTFFIAFGAAIAYGLYLVRRNGSIDPDATAADVPAGT 292
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 417 IFIGKGFVFGILGIAVAVILSAIGTYFFGYKDEMDAGI-----APT 457

QY 293 TKAEAEAPAEFSNDSTIIQAPLTGAIALSSVSDAMFASGKLGSGVAIVPTKGQLVSPVS 352
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 458 TKEAKETGE-----AEVIVSPTRGNIVFLNEVKDEAFAGSLGKGVAVIPOPGLISPVN 513

QY 353 GKIVAFPSHAFVTRKAEDGSNDIILMHIGFTVNLNGTHFNPLKKQGEVKAAGELLIC 412
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 514 GTIETAFTHGAIGRS-----DKGVEILLHVGFTVQLNGKYFKLVVAGQORVLVQALL 569

QY 413 BFDIDAIIKAAGYEVTPPIVSVN 434
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 570 BFDLEAIKADGVDTTPIVSVN 591

RESULT 4
US-10-474-776-375
; Sequence 375, Application US/10474776
; Publication No. US20040110181A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; TITLE OF INVENTION: NOVEL STREPTOCOCCUS PNEUMONIAE OPEN READING FRAMES ENCODING POLYH
; TITLE OF INVENTION: ANTIGENS AND USES THEREOF
; FILE REFERENCE: AM100649-PCT
; CURRENT APPLICATION NUMBER: US/10/474,776
; CURRENT FILING DATE: 2003-10-14
; NUMBER OF SEQ ID NOS: 752
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 375
; LENGTH: 655
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-474-776-375

Query Match 28.5%; Score 673; DB 16; Length 655;
Best Local Similarity 35.3%; Pred. No. 2.8e-49;
Matches 170; Conservative 87; Mismatches 173; Indels 52; Gaps 16;

QY 1 MAMVFPISLVNGYDVAATWAGEMP--MWSLFGLDVAQAGYQGTVLPLVVSVILATIEKF 58
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 195 ICLVSPOLLNAYAVATPAADIAANWNNFGYVTNRIGYQAQVIPALLAGLSLYLEIF 254

QY 59 LHKRLKGTADFLITPLVLLLTGFTTFAIGPAMRWGDLVLAHGLQGLYDFG--GPVG-- 114
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 255 WKHHIPEVISMIFPFLSLPALLIAHTVLGP-----IGWTIGQLSSVVLVAGLTGPVKWL 310

QY 115 -GLLFGVIVSPVITGLHQSPFIELFL-NQGSFIFATASMANIAQGAACLAFFFLAK 172
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 311 FGAIFGALYAPFVITGLHHTNAIDTQLIADAGTALWPMIALSNIAQGSVAFFYFMR 370

QY 173 -SEKIXGLAGASGVSALVIGITEPAIFGNLRWPFRTIGTAAIGGALIALENIKAV 231
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 371 HDBREAQVSLPATISAYLIGTEPALFGVNVKYYIYFFVAGMTGSALAGMLSVTFNVTAASI 430

QY 232 GAAGFLGWSIDAPDMVME-----LVCAVVTFFIAFGAAIAYGLYLVRRNGSIDPDATAAP 287
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 431 GIGGLFGILSIQPMPLFPAGTMLVAVVPMPLTFF-----FFRKAGLFTK----- 475
```

```
QY 288 VPAGTTKAEAE-----APAEFSNDSTI-----IQAPLTGEATIALSSVSDAMFASGKLGSSG 337
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 476 -TEGDTNQAEEVAQEEAEFVNHEPVELTSVEIISPLTGQVKELSOATDPIFASGVMGQG 534

QY 338 VAIVPTKGQLVSPVSGKIVVAFPSGHAFVTRKAEDGSNDIILMHIGFTVNLNGTHFN 397
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 535 LVIEPSQELTSPVNGTIVLFTPKHAGI--VSDG--VELLIHIGMDTVGLDGKGFES 590

QY 398 LKKQDEVKAGELLCEFPIDAIKAAGYEVTPPIVSVN---YKKTGPVNTYGLGEIEAGAN 454
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 591 LVVQGDHVTVGQQLIRFDMVKAAGLVTEFTPIITNQDAYTATIP-GTYPT-TIOAGAS 648

QY 455 IL 456
   : :
Db 649 LM 650

RESULT 5
US-09-815-242-13467
; Sequence 13467, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13467
; LENGTH: 627
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-815-242-13467

Query Match 27.9%; Score 660; DB 9; Length 627;
Best Local Similarity 34.3%; Pred. No. 3.5e-48;
Matches 159; Conservative 92; Mismatches 187; Indels 26; Gaps 10;

QY 1 MAMVFPISLVNGYDVAATWAGEMPMSLFGLDVAQAGYQGTVLPLVVSVILATIEKFLH 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 182 MMLVSGSLPNAWAVA---OGGEVTAMNFFGF-IPVVGLQGSVLPAFTIGVVGAKFEKAVR 237

QY 61 KRLKGTADFLITPLVLLLTGFTTFAIGPAMRWGDLVLAHGLQGLYDFGPGVGLLFG 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 238 KVPDVIDLLVTFEVLVWMSILGLFVIGVFHVVENYILIAKILMSFPGGLGFLIG 297

QY 121 YSPVITGLHQSPFIELFLFNQGSFTF-ATASMANIAQGAACLAFFFLAKSEKLG 179
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 298 VHQLIVSVGVHHTFNLLLEVQLLAADHANPENAIITAAWTAQGAATVAVGVKTKNPKLKL 357
```



```
QY 180 AGASGVSAVLGITEPAIFGVNLRWRPFIFIGTAAIGGALIALFNKAVALCAAGFLG- 238
Db 358 AFAALSAFLGITEPAIFGVNLRWRPFIFISLAGAIGGLASILGAGTGNGITTIPTG 417
QY 239 VVSIDAPDMVFMFLVCVAVTFFIAFGAAIAYGLYVVRNGSIDPDATAAPVPAGTTKAEAE 298
Db 418 MLVYVNGQLPOYLLMVAVSFALGFALTYMGY-----EDEVDAATAAKQAEVAEKEE 470
QY 299 -APAEFSDSTIIQAPLTGSAIALSSVSDAMFASGKLGSGVAIVPTKGQVSPVSGKIVV 357
Db 471 VAPAAALQNETLV--TPIVGDVALADVNDVFSSGAMGQGIAPKPSQGVVYALADAEVSI 538
QY 358 APSPGHAFVTRKAEDGNSVDILMHIGFDVNLNGTHFNPLKKGQDEVKAGELLCEBFDID 417
Db 529 APTGHAFGLKTR-----NGAEVLHVIGDITVSMNGDGFKAQVQGNKVRAGDVLGTGDSN 584
QY 418 AIKAAAGVEVTTPIVNSN---YKKTGPVNTYGLGEIEAGANLLNV 458
Db 585 KIAAAGLDDTTWIVNTADYASVAPVAT---GSVSKGDVAVIEV 625

RESULT 6
US-10-474-776-367
; Sequence 367, Application US/10474776
; Publication No. US20040110181A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; TITLE OF INVENTION: NOVEL STREPTOCOCCUS PNEUMONIAE OPEN READING FRAMES ENCODING POLY
; FILE REFERENCE: AM100649-PCT
; CURRENT APPLICATION NUMBER: US/10/474,776
; CURRENT FILING DATE: 2003-10-14
; NUMBER OF SEQ ID NOS: 752
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 367
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-474-776-367

Query Match 24.8%; Score 585; DB 16; Length 381;
Best Local Similarity 34.4%; Pred. No. 5.4e-42;
Matches 136; Conservative 82; Mismatches 155; Indels 22; Gaps 8;

QY 70 LIPTVILLTGTFTTAIGPAMRWGVDLHAGLQGLYDFGPGVGGILFGLVYSPVITG 129
Db 1 MVTPEVTLVMSILGLFVIGFVHVHVENYIILIAIKAILSMPPFGLGGFLIGVHQLIVSG 60
QY 130 LHQSFPPPELEFNQGGSFIF-ATASMANIAQGAACLAFFFLAKSEKLKGLAGASVSAY 188
Db 61 VHHLENLEVLQLLAADHANPNALITAMTAQGAATVAVGVKTKNPKLKTAFPAALSAP 120
QY 189 LGITEPAIFGVNLRWRPFIFIGTAAIGGALIALFNKAVALCAAGFLG-VVSIDAPDM 247
Db 121 LGITEPAIFGVNLRWRPFIFISLAGAIGGLASILGAGTGNGITTIPTGTMLYVNGQL 180
QY 248 VMLVCVAVTFFIAFGAAIAYGLYVVRNGSIDPDATAAPVPAGTTKAEAE-APAEFSD 306
Db 181 POYLLMVAVSFALGFALTYMGY-----EDEVDAATAAKRAEVAEKEEVAPALQNE 233
QY 307 STIIQAPLTGSAIALSSVSDAMFASGKLGSGVAIVPTKGQVSPVSGKIVVAPFSGHAF 366
Db 234 TLV--TPIVGDVALADVNDVFSSGAMGQGIAPKPSQGVVYAPADAESAIFPTGHAFG 291
QY 367 VTRKAEDGNSVDILMHIGFDVNLNGTHFNPLKKGQDEVKAGELLCEBFDIDAIKAAGVEV 426
Db 292 LKTR-----NGAEVLHVIGDITVSMNGDGFETKVAQGNKVRAGDVLGTGDSNKKIAAAGLDD 347
QY 427 TPIVNSN---YKKTGPVNTYGLGEIEAGANLLNV 458
Db 348 TTMVINTGDTYASVAPVAT---GSVAKGDVAVIEV 379
```

```
RESULT 7
US-10-282-122A-73843
; Sequence 73843, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR FILING DATE: 2003-02-20
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-05-23
; PRIOR FILING DATE: 2000-05-23
; PRIOR FILING DATE: 2000-05-26
; PRIOR FILING DATE: 2000-09-06
; PRIOR FILING DATE: 2000-11-27
; PRIOR FILING DATE: 2000-12-22
; PRIOR FILING DATE: 2000-12-22
; PRIOR FILING DATE: 2001-02-09
; PRIOR FILING DATE: 2001-02-09
; PRIOR FILING DATE: 2001-02-16
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 73843
; LENGTH: 612
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-282-122A-73843

Query Match 24.5%; Score 578.5; DB 15; Length 612;
Best Local Similarity 33.3%; Pred. No. 3.8e-41;
Matches 143; Conservative 83; Mismatches 143; Indels 61; Gaps 12;

QY 27 SLFGLDVQAQYQGTVPVLVSVMLATIEFKLHKLKGTADFLITPVLTLLTGLTFI 86
Db 196 SLFGVPVTPASVGSVVPFIIMVLMKYIEKMIKLTFAVTKSLFQPLTVLLVSSCIALV 255
QY 87 AIGPAMRWGVDLHAGLQGLYDFGPGVGGILFGLVYSPVITGHLQSPPELEFNQGG 146
Db 256 VVGPIGVVBGGLNVLQGMVDVAGWLTALILGALIMPVMTGHWAFAP----- 305
QY 147 SFIFATASMA-----NIAQGAACLAFFFLAKSEKLKGLAGASVSAYL-GIT 192
Db 306 --IFLAASIAFPDVLILFAMLGSLNLAQGAASMAVALKSNNTKQIAFAAGFSALLAGIT 363
QY 193 EPAPFGVNLWRWPFIFIGTAAIGGALIALF-----NIKAVALGAAGFLGV---VSI 245
Db 364 EPALYGVTLKPKPLY-----AAMTGGGAGLGFAGITSVKAYLFAVPSLIALPQFTYSDVP 419
QY 246 -DMWMFLVCVAVTFFIAFGAAIAYGLYVVRNGSIDPDATAAPVPAGTTKAEAEAPAEFS 304
Db 420 SNVNALIVANISVITFVLAYIFG-----IDEESS-----SNLEVKAGVS 461
```


Db 462 SQVKSGLSTKQI-LYAPMTGEMFLSEVDETFSSKLLGEGFALLPSEGEVYAPDFGEVI 520
QY 357 VAPPGHAFARTKADGNSVDILMIHGDTVNILNTHNPLKQGDVYKAGELLCEFDI 416
Db 521 TFPPTKHAVALK-----NTRGVEVLHVGIDTVELKQGFQOLVSVGDVVKRGQALLKMDI 576
QY 417 DAIKAAGYEVTPPIVYSN 434
Db 577 DFITSXGYSLISEVVVTN 594

RESULT 10
US-10-282-122A-70209
; Sequence 70209, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR FILING DATE: 2003-02-20
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 70209
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-282-122A-70209

Query Match 20.8%; Score 492; DB 15; Length 379;
Best Local Similarity 37.3%; Pred. No. 5.9e-34;
Matches 100; Conservative 63; Mismatches 89; Indels 16; Gaps 5;
QY 1 MAMVFPVSLVNGVDVATMAAGEMPMWSLFGLDVAQAGYQGTGTVLPVLVSVSWILATIEKFLH 60
Db 98 LILMHPQLVSQXD-----LAKGNIPWNLFGLBIKQLNQYQQVLPVLIAYVLAKIEKGLN 153
QY 61 KKLKGTADPLTTPVLTLLITGLTFTTAIGPAMRWCDVLAHGLQGLYDFGPGVGLLFL 120
Db 154 KVVHDSIKMLVGPVALLVTGFLAIIIGVALLIGTGITSGVTTFIQHAGMIGGAIYGL 213

QY 121 VYSPVITGLHOSFPPIELELNFQ--GGSFIFATASMANIAQGAACLAIVFFLAKSEKL-- 176
Db 214 LYAPLVITGLHMFADVDFQMLGSSLGTYLWPIVAISNICGSAAFGAWFYVKKRWVK 273
QY 177 -KGLAGASVSAVLGTETPAIGVNLRLRWPFIFIGTAAIGGALIALFNKAVAGAAAG 235
Db 274 BEGLALTSCISGLMGLVTEPAMFGVNLPLKYPFTTAAISTSCVLGAIYGMNVN----LKGVG 329
QY 236 FLGV--VSDAPDMVMFLVCAVVTFFI 260
Db 330 VGVPAFISIQKEFPVYLIATAIAIV 357

RESULT 11
US-09-738-626-6961
; Sequence 6961, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIALI, KEIKO
; APPLICANT: YOKOL, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENO, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6961
; LENGTH: 683
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6961

Query Match 19.8%; Score 468.5; DB 9; Length 683;
Best Local Similarity 29.8%; Pred. No. 1.5e-31;
Matches 141; Conservative 79; Mismatches 206; Indels 47; Gaps 9;
QY 27 SLFGLDVAQAGYQGTGTVLPVLVSVSWILATIEKFLHKLKGTADFLITPVLTLGLLGLTFI 86
Db 218 TVFGLPMVLYNDYSGVFPPLIAAIGLYVWVEXGLKIIPEAVQMVVPVFPFSLIMIPATAF 277
QY 87 AIGPAMRWGDVLAHGLQGLYDFGPGVGLFGLVYSIVITGLHOSFPPIELELNFQ-G 145
Db 278 LLGPPFGICVNGISNLEAINNFPFSLIVIPLLYPLVPLGLHWPINAIMQINILG 337
QY 146 GSFIFATASMANIAQGAACLAIVFFLAKSEKLKLAGAS--GVSAVL--GITEPAIFGVNL 201
Db 338 YDFIQGPMGAWNFACFGLVTVGVLLSIKERNKAMQVSLGGMLAGLGGISEPSLYGVLL 397
QY 202 RLRFPPFIFIGTAAITGALIALFNKAVAGAGFLGVVSDAPDMVMFLVCAVVTFFIA 261
Db 398 RPKTYFRLLPGLAGGIVMGIFDIKAYAFVFTSLITIPAMD--PWLGYTIGIAVAFVVS 455
QY 262 FGAAATAGLYLVRRNGSID-----PDAT-AAAPV----- 288
Db 456 MFLVALD---YRSNEERDEAPAKVAQKQAEEDLKAEANATPAAPVAAAGAGAGAGA 512
QY 289 PAGTTTAKAEAPAEFNSDSTIIQAPLTGEAIALSSVSDAMFASGKLGSGVAVLPTKGQV 348
Db 513 AGAATAVAAKFLAAGEVWDIVSFLGKAIPLSEVDPDPIFAAGKLGPGIAIQPTGNTVV 572


```
QY 309 IIOAPLTGEAIALSSVSDAMFASGKLGSGVAIVPTKQGLVSPVSGKIIVAFPSGHAFVAVR 368
Db 534 IIVHAPLTGEVTPLESEVPDQVFSKMMGDDGIAIKPSQGEVRAPFNGKIQMIFPTKHAIGL- 592
QY 369 TKAEDGSNDIILMHIGFDVTNLTNGTHFNPLKKGDEVKAGELICEFDIDAIAKAAGYEVTT 428
Db 593 ---VSDSGLELLIHIGLDTVKLNGEGFTLHVESGQEVKQGDLLINFDLDYIRNHAUSDIT 649
QY 429 PIVVS 433
Db 650 PIVVT 654

RESULT 14
US-09-815-242-5658
; Sequence 5658, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5658
; LENGTH: 679
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5658

Query Match 14.6%; Score 345.5; DB 9; Length 679;
Best Local Similarity 24.2%; Pred. No. 6.3e-21;
Matches 132; Conservative 68; Mismatches 192; Indels 153; Gaps 17;

QY 27 SLFLGDLVAQAGYGGTFLVPLVSVWTLATIEKFLHKLKGTADFL----ITPVLTLTLTG 82
Db 123 SILGIPTLTGTGTVFGGIIIGALAAW---CVNKEYNINLPSYLGFFAGKRFVPIM-MATTSF 178
QY 83 LTFPAIGAMRWGVDLAHLQ----GLYDFGPGVGGILFGLVYSPVITGLHQSF--- 134
Db 179 I--LAFPMALIW--PTIQTGLNAFTGLDSTNGTAVFLGFIKRLIPFGHLHIFHAPF 234
QY 135 -----PPIBLELNFQG-----GSFIPATASMANIAQGAACLAFFFLA 171
Db 235 WFEFGSKNAAGEIIGHGDIQIFIEQIREGAHLTAGKFMQGEPPVMFGLPAAALAIYHSA 294
QY 172 KSEKLGKAGAGVSA-----VLGITPEPAIFGNLRLWFFFGIGTAAIGGALIAFNK 227
Db 295 KPNKKVWAGLMSAALTSFLTGITPEPFSFLFVAPLLFFI---HAYLDGLSFUTLYLL 351
```

```
QY 228 AVALG---AAGFLGVVSIDA-PDMVMF-----LVCVVVTFPI----- 260
Db 352 DVHLGYTFSGGFDYVLLGLVLPNKQWLIVPGLVAVIYVYFVFRFLIVKLYKTPGRE 411
QY 261 -----AFGAIAIAGLY-----LVR----- 274
Db 412 DKSOQAVTASATELPYAVLEAMGGKANI KHLDACITRLRVEVNDKSKVDVPGKLDLGASG 471
QY 275 -----RNGSIDPDATAAPVPAGTTTKAEAEAPAEFNSDST 308
Db 472 VLEVGNNMQAIFGPKSDQIKHEMQIMNGOVVENPTTMEDDKDETIVVAEDKSATSELSH 531
QY 309 IIOAPLTGEAIALSSVSDAMFASGKLGSGVAIVPTKQGLVSPVSGKIIVAFPSGHAFVAVR 368
Db 532 IIVHAPLTGEVTPLESEVPDQVFSKMMGDDGIAIKPSQGEVRAPFNGKIQMIFPTKHAIGL- 590
QY 369 TKAEDGSNDIILMHIGFDVTNLTNGTHFNPLKKGDEVKAGELICEFDIDAIAKAAGYEVTT 428
Db 591 ---VSDSGLELLIHIGLDTVKLNGEGFTLHVESGQEVKQGDLLINFDLDYIRNHAUSDIT 647
QY 429 PIVVS 433
Db 648 PIVVT 652

RESULT 15
US-09-815-242-12270
; Sequence 12270, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12270
; LENGTH: 681
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12270

Query Match 14.6%; Score 345.5; DB 9; Length 681;
Best Local Similarity 24.2%; Pred. No. 6.4e-21;
Matches 132; Conservative 68; Mismatches 192; Indels 153; Gaps 17;

QY 27 SLFLGDLVAQAGYGGTFLVPLVSVWTLATIEKFLHKLKGTADFL----ITPVLTLTLTG 82
Db 125 SILGIPTLTGTGTVFGGIIIGALAAW---CVNKEYNINLPSYLGFFAGKRFVPIM-MATTSF 180
```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 29, 2004, 23:06:04 ; Search time 39 Seconds
(without alignments)
795.816 Million cell updates/sec

Title: US-09-604-231-2

Perfect score: 2363

Sequence: 1 MAMVPSLVNGYDVAAATWRA.....IEAGANLNLNVAKEAVPAP 468

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/1aa/5A-COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B-COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A-COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B-COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PCUS-COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	832.5	35.2	496	US-09-134-000C-5822	Sequence 5822, Ap
2	669	28.3	627	US-09-583-110-3124	Sequence 3124, Ap
3	662	28.0	628	US-09-107-532A-5288	Sequence 5288, Ap
4	655	27.7	655	US-09-583-110-2958	Sequence 2958, Ap
5	626.5	26.5	656	US-09-489-039A-8212	Sequence 8212, Ap
6	609	25.8	670	US-09-134-000C-4606	Sequence 4606, Ap
7	583.5	24.7	583	US-09-107-532A-6811	Sequence 6811, Ap
8	572.5	24.2	612	US-09-583-110-2911	Sequence 2911, Ap
9	569	24.1	634	US-09-543-681A-7019	Sequence 7019, Ap
10	562.5	23.8	631	US-09-107-532A-3902	Sequence 3902, Ap
11	537	22.7	316	US-09-134-000C-5482	Sequence 5482, Ap
12	520.5	22.0	243	US-09-134-001C-4114	Sequence 4114, Ap
13	501.5	21.2	427	US-09-107-532A-5530	Sequence 5530, Ap
14	497.5	21.1	526	US-09-134-000C-4715	Sequence 4715, Ap
15	489	20.7	483	US-09-489-039A-13018	Sequence 13018, A
16	412	17.4	590	US-09-107-532A-6004	Sequence 6004, Ap
17	363.5	15.4	481	US-09-543-681A-5156	Sequence 5156, Ap
18	357.5	15.1	478	US-09-489-039A-8977	Sequence 8977, Ap
19	336.5	14.2	478	US-09-134-001C-5065	Sequence 5065, Ap
20	302.5	12.8	696	US-09-134-001C-4296	Sequence 4296, Ap
21	295	12.5	726	US-09-583-110-2782	Sequence 2782, Ap
22	294.5	12.5	635	US-09-489-039A-13395	Sequence 13395, A
23	291	12.3	656	US-09-107-532A-5083	Sequence 5083, Ap
24	290	12.3	688	US-09-543-681A-5236	Sequence 5236, Ap
25	284	12.0	482	US-09-107-532A-6644	Sequence 6644, Ap
26	283	12.0	585	US-09-107-532A-4366	Sequence 4366, Ap
27	275	11.6	482	US-09-489-039A-9909	Sequence 9909, Ap

28 269.5 11.4 460 4 US-09-543-681A-5773 Sequence 5773, Ap
29 267 11.3 170 3 US-09-134-001C-3134 Sequence 3134, Ap
30 266.5 11.3 470 4 US-09-107-532A-5798 Sequence 5798, Ap
31 260 11.0 527 4 US-09-634-238-368 Sequence 368, Ap
32 256 10.8 196 4 US-09-543-681A-7659 Sequence 7659, Ap
33 248 10.5 315 4 US-09-134-000C-3657 Sequence 3657, Ap
34 244 10.3 172 4 US-09-489-039A-9162 Sequence 9162, Ap
35 230 9.7 686 4 US-09-252-991A-18115 Sequence 18115, A
36 223 9.4 479 4 US-09-489-039A-10571 Sequence 10571, A
37 181.5 7.7 153 4 US-09-107-532A-4365 Sequence 4365, Ap
38 159 6.7 355 4 US-09-107-532A-5592 Sequence 5592, Ap
39 157 6.6 448 4 US-09-583-110-4040 Sequence 4040, Ap
40 151.5 6.4 658 4 US-09-492-709A-352 Sequence 352, Ap
41 144 6.1 101 4 US-09-134-000C-3693 Sequence 3693, Ap
42 141.5 6.0 650 4 US-09-583-110-5074 Sequence 5074, Ap
43 139.5 5.9 451 4 US-09-107-532A-5352 Sequence 5352, Ap
44 137 5.8 495 4 US-09-252-991A-19278 Sequence 19278, A
45 136.5 5.8 661 3 US-09-134-001C-4303 Sequence 4303, Ap

ALIGNMENTS

RESULT 1

US-09-134-000C-5822
; Sequence 5822, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5822
; LENGTH: 496
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-5822

Query Match 35.2%; Score 832.5; DB 4; Length 496;
Best Local Similarity 39.3%; Pred. No. 2.4e-75;
Matches 184; Conservative 84; Mismatches 171; Indels 29; Gaps 7;
QY 1 MAMVPSLVNGYDVAAATWRAAGEMPMSLFLGLDVAQAGYQGTVPVLVTVSVILATIEKFLH 60
Db 47 MIMTPDJ-----GGATEYWNIFGVHVAQTNVAYQVIVLASVYLLILEKTFH 95
QY 61 KRLKGTADFLITPVLTLLTGLTFATIGPANRWGDVLAHGLQGLYDFGPGVGLLFLGL 120
Db 96 KKLPSIDFTFTPLLSVITIGLFTFTVGPMLLSNGITDAIWLINATGFIGMGIFGG 155
QY 121 VYSPITVGLHQSPPIELFLNQ-----CGSFIFATASMANIAQGAACLAFFLAKSE 174
Db 156 TYSLIVMTGLHQSPPIETQLLSAWTNGIGHGDFIVVASMANVAQGAATAIWFLTKNS 215
QY 175 KLKGIAGASGVSVLGITETPAIFGVNLRWPPFPGTAAIGGALIALFNKVALGAA 234
Db 216 KTKSLASAGLSALGITEPALFGVNLKYRFPFPFCALIGSGIAAAITGLLVAVSLGSA 275
QY 235 GFLGWSIDAPDMVWFLVCAVWTFPIAFGAIAAGLVYVRENGSIDDPATAAPVPAGTK 294
Db 276 GFLGFLSINATSIFFYLLCELISVFTAYFYG--RTRSSSIFAAEALAEQTSVDTSR 333
QY 295 AEAAPAEFNDSTI--IQAPLTGEALISVSDAMFASGKLGSVATVPTKGLQVSPVS 352
Db 334 INTNQIANADEPTTETIVISPLAGEITTLGSVNDPVSFSSBSIGKGIKPNGTIYSPVD 393
QY 353 GKIIVAPSPGHFAVTRKADGGSNVDILMHIGFTVNLNGTHFNPLKKQG--DEVKAGEL 410

Db 394 GIVQVVFETGHAYDLKS----NTGABLLIHVGIDTVSLNGKGT--KKVGAQKVKKEGV 447
QY 411 LCFEDIDAIKAAGYEVTPPIVWSNYKKTGPNVNYGLGEIBAGANLNV 458
Db 448 LGTFDSVTIINSGLDTTWVITNSKDYSEVIPITKNIIVTEGAALLTI 495

RESULT 2
US-09-583-110-3124
; Sequence 3124, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 3124
; LENGTH: 627
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-3124

Query Match 28.3%; Score 669; DB 4; Length 627;
Best Local Similarity 34.5%; Pred. No. 1.2e-58;
Matches 160; Conservative 92; Mismatches 186; Indels 26; Gaps 10;

QY 1 MAMVPPSLVNGYDVAATMAAGEMPWMSLFLGLDVAQAGYQGTVPVLVWVSVILATIEKFLH 60
Db 182 MNLVSGSLPNAWAVA---QGEVITAMNFFGP-IPVVGLOGSVLPAPFIIGVGAKEKAVR 237
QY 61 KRLKGTADFLITPVLTLTLTGLTFIAIGPAMRWGVLDVLAHGLQGLDFGPGVGLLFLGL 120
Db 238 KVPDVIDLLVTPFTVLLVMSILGLFVIGPVHVVENVILIATKAILSIPFLGLGFLGG 297
QY 121 VYSPVITGLHQSFPPIELHFNQGGSPIF-ATASMANIAQGAACLAIVFFLAKSEKLG 179
Db 298 VHQLIWSGVHIFNLLEVQLLAADHANFPFNAIITAAMTAQGAATVAVGVTKNPKLKITL 357
QY 180 AGAGSVSAVLGITEPAIFGVNRLRWPFFIGTAAIGGALIALFNKAVALGAAGFLG- 238
Db 358 APPAALSAPLGLTEPAIFGVNRLFRKPFPLSLIAGAGGLASILGLAGTNGNITIIPT 417
QY 239 VVSDIDAPDMWFLVCAVVTFFIAFGAATAIYGLYVVRNGSIDPDATAPVPAGTTKARAE 298
Db 418 MLYVNGQLPQYLLWAVSFALGFALTVMFGY-----EDEVDATAAAKQAEVAEKEE 470
QY 299 -APAFNSDSTIIQAPLTGEATALSVDAMFASGKLSGVVAIVPKGOLVSPVSKIVV 357
Db 471 VAPAAALONETLV--TPIVGDVVALADVNDPVFSSGAMGGIYAVKPSQGVVYAPADAESVI 528
QY 358 APFSGHAFVRUKAEDGNSVDIIMHIGFTVNLNGHFNPLKKQGDGEVKAAGELLCEFDID 417
Db 529 APPTGHAFGLKTR----NGAEVLIHVGIDTVSMNGDGFKAQVQGNKVXAGDVLGTFDSN 584
QY 418 AIIKAAGYEVTPPIVWSN---YKKTGPNVNYGLGEIBAGANLNV 458
Db 585 KIAAGLDTTWVITNTADYASVAPVAT---GSVAKGDVAVIEV 625

RESULT 3
US-09-107-532A-5288
; Sequence 5288, Application US/09107532A
; Patent No. 6583275

GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS: 7310
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 5288:
SEQUENCE CHARACTERISTICS:
LENGTH: 628 amino acids
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...628
SEQUENCE DESCRIPTION: SEQ ID NO: 5288:
US-09-107-532A-5288

Query Match 28.0%; Score 662; DB 4; Length 628;
Best Local Similarity 33.5%; Pred. No. 6.1e-58;
Matches 160; Conservative 89; Mismatches 172; Indels 56; Gaps 10;

QY 1 MAMVPPSLVNGYDVAATMAAGEMPWMSLFLGLDV--AQAGYQGTVPVLVWVSVILATIEKF 58
Db 185 MALVVP-----AITAVAGAGEAISFFGIPVILSPSYTSSVIPILAVWVQSKLEPF 236
QY 59 LHKRLKGTADFLITPVLTLTLTGLTFIAIGPAMRWGVLDVLAHGLQGL-----YDFGGPVG 114
Db 237 VKKVIPOQLMQLLVLLVWVWVPLTFLALGP-----IGTVAGNALGGLFNSIYGFSPIVA 292
QY 115 GLLFGLVYSPVITGLHQSFPPIE-LELFNQGGSFIFATASMANIAQGAACLAIVFFLAKS 173
Db 293 GLIMGSLMQVFMFGMHGQFVPMFLNTEQYGFVLMPLLPAILLAQGAALAVALTGD 352
QY 174 EKLKGLAGASGVSAVLGITEPAIFGVNRLRWPFFIGTAAIGGALIALFNKAVA--- 230
Db 353 TKLRALGISTVTSLSFGITEPVYGVTLPLKKPFIAACISGGIGGAILGFSGVKAPSSSL 412
QY 231 ---LGAAGFLGVSDIDAPDMWFLVCAVVTFFIAFGAATAIYGLYVVRNGSIDPDATAP 287
Db 413 VSLITPTPTINTVDGVESNVTVAIVATGIAFVLTILGFDQEQTNQLE----- 465
QY 288 VPAGITKAAEAPAFNSDSTIIQAPLTGEATALSVDAMFASGKLSGVVAIVPKGQL 347

Db 466 -----NKHANAGEPITSARHTLKSPLTGKVLSEVPDQVSSGVMGKGIADPEVGE 519
QY 348 VSPVSGKIVVAFPSGHAFVTRKAEDGNSVDILMHIGFDTVNLNGTHENPLKKGDEVKA 407
Db 520 VAPADGEITITFTGHAVGITT--TDGA--EILIHGMDTVELNGCFEILVKQGLVKA 575
QY 408 GELLCEFDIDAIAKAGYEVTTPIVSNYKTKGPVNTYGLGEIBAGANLLNVAKEAV 464
Db 576 GDLIRFDIEAIRAAGSVITPVITN-----TDAFADILELDQKEII 618

RESULT 4

US-09-583-110-2958
; Sequence 2958, Application US/09583110
; Patent No. 669703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 2958
; LENGTH: 655
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-2958

Query Match 27.7%; Score 655; DB 4; Length 655;
Best Local Similarity 34.4%; Pred. No. 3.4e-57;
Matches 165; Conservative 91; Mismatches 177; Indels 46; Gaps 15;
QY 1 MAWFFPSLVNGYDVAATMAAGEMP--MWSLFLDVAQAQYQGTVLPLVNVSWILATIEKF 58
Db 195 ICLVSPQLLNAYAVASTAADAANWNVNFGYFTVNRIGYQAVIPALLAGLSLSYEIF 254
QY 59 LHKRLKGTADFLITPVLTLTLTGFTTFAIGPAMRWGVDVLAHGLQGLYDFG--GPVG-- 114
Db 255 WRKHIPVISMIFVFPFLSLIPALILAHTVLGP-----IGWTIGQLSSVVLAGITGPVKWL 310
QY 115 -GLLFGLYSPITVTLHQSPFPIELF--NOGGSFIFATASMANIAQGAACLAFFFLAK 172
Db 311 FGAIFGALYAPVITGLHHTNAIDTQLIADAGGTALWPMIALSNIAQGSVAFFYFMHR 370
QY 173 SEKLGK-LAGASGSVAVLGITEPAIFGVNLRWRPFFIGTAAIGGALLIALFNKAVAL 231
Db 371 HDECEAQLSPATISAYLGVTPEALFGVNVKIIYFVAGMTGSALAGMLSVTFNTAASI 430
QY 232 CAAGFLGVVSDADPMVMF-----LVCVVVTFPIAFGAALAYGLYLVRNG--SIDPAT 284
Db 431 GIGLPGILSIQPOWMLPFAGTMLVAIVVPLLTP-----PFRKAGLFTKLEGTN 481
QY 285 AAPVPAGTTKAE--AEAPBSNDSTIIQAPLTGEAIALSSVSDAMFASGLSGVAIVTKGQL 342
Db 482 LQABFVAQEEAEFVSHEPVELT--SVEIISPLTGQVKLSQATDPVFASGVMGQGLVIEP 539
QY 343 TKGQLVSPVSGKIVVAFPSGHAFVTRKAEDGNSVDILMHIGFDTVNLNGTHENPLKKG 402
Db 540 SQGLTSFVNTVTVLFTKKAIGT--VSDG--VELLIHGMDFVLDGKGFSLVVQG 595
QY 403 DEVKAGELLCFFDIDAIAKAGYEVTTPIVSNYK--TGPVNTYGLGEIBAGANLL 456
Db 596 DHVIVGQQLIRFDMVKAAGLVETPVIINQDAYTATITGTPT-----TIQAGASLM 650

RESULT 5

US-09-489-039A-8212
; Sequence 8212, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 8212
; LENGTH: 656
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-8212
Query Match 26.5%; Score 626.5; DB 4; Length 656;
Best Local Similarity 33.5%; Pred. No. 2.6e-54;
Matches 158; Conservative 82; Mismatches 183; Indels 49; Gaps 11;
QY 2 AMVFPFSLVNGYDVAATMAAGEMPMSLFLDVAQAQYQGTVLPLVNVSWILATIEKFLHK 61
Db 219 ALTHPLMTQAPE--ASQAPG--NAVEHFLGIPVTFINYSVSPVPIIILASVWSCWLERKSA 275
QY 62 RUKGTADFLITPVLTLTLTGFTTFAIGPAMRWGVDVLAHGLQGLYDFGPGVGGLLFGLV 121
Db 276 LLPSSMKNFPSAICLAVVVPVTLFLVIGPVATWLSHLLANGYQFIYAFAPWLAGAVLGAM 335
QY 122 YSPIVITGLHQSPFPIEL--ELFNQGSFIFATASMANIAQGAACLAFFFLAKSEKLGLA 180
Db 336 WQCVIFGLHGLVFLMNMNMTVLGHSMLPTILPAVIAQVAGVLGIFLATRDARQVLA 395
QY 181 GASGSVAVLGITEPAIFGVNLRWRPFFIGTAAIGGALLIALFNKAVAG-----AA 234
Db 396 GSAPGAGLFGITEPAIYGLTLPLRPPFFGCVAGAIQGAITAFSNSYAYSGLENIFFPA 455
QY 235 GFLGVVSDADP-----DMVMFLVCVVTFPIAFGAALAYGLYLVRNGSIDPDATAAP 287
Db 456 QMIPPGGIDASVWGGLIGTGVAFLACVLTFPA-----GL-----PRGSAAP 497
QY 288 VPAGTTKAEAPAFBSNDSTIIQAPLTGEAIALSSVSDAMFASGLSGVAIVTKGQL 347
Db 498 -----GAVTVAPAS-AND---ILAPMSGVIALEQVDPDSTFASGLLGKGVAIIPAVGQV 547
QY 348 VSPVSGKIVVAFPSGHAFVTRKAEDGNSVDILMHIGFDTVNLNGTHENPLKKGDEVKA 407
Db 548 IAPFGEVASLFOYKHAIGL-----QSDSGIELLIHVGDITVKLDGVPPTAHVKGSDRVOA 603
QY 408 GELLCEFDIDAIAKAGYEVTTPIVSNYKTKGPVNTYGLGEIBAGANLLNVA 459
Db 604 GDLLIEFDRAILDAGYDLVTPIIISNDDYREIDTVASSAVEAGQPLLSVS 655

RESULT 6

US-09-134-000C-4606
; Sequence 4606, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4606
; LENGTH: 670


```
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 2911
; LENGTH: 612
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-2911

Query Match      24.3%; Score 572.5; DB 4; Length 612;
Best Local Similarity 33.3%; Pred. No. 6.7e-49;
Matches 143; Conservative 81; Mismatches 145; Indels 61; Gaps 12;

QY 27 SLFGLDVAQAGYOGTVLPVLVSVWILATIEKFLHKLKGTADFLIIPVLTLITLGTFTFI 86
Db 196 SLFGAPVTASYGSSVVPILIMVLMKYTEKMTAKLTPAVTKSFLOFTLVLVSSCIAIV 255

QY 87 AIGPAMRWGVDVLAHGLQGLYDFGPGVGLLFGLVYSPVITGLHQSFPPIELFLNQGG 146
Db 256 VVGPFGIVGEGLSNLVGOMYGAGWLTALILGAIIMPFIWTHWAFAP----- 305

QY 147 SFIFATASMA-----NTAAGAACLAFFFLAKSEKLGAGAGSVSAVL-GIT 192
Db 306 --IFLAASIATPDVLILPAMLGSLNLAQGAASMAVALKSKNNNTKQIAFAAGFSALLAGIT 363

QY 193 EPAIFGNLRLRWPFFIGTGTAAIGGALIALF-----NIKAVAGAGFLGV---VSIDAP 245
Db 364 EPALGVTLKTKPLF---AAMIGGGLAGLFGAGTSVKAYLFAVPSLIALPQFIYSYDVP 419

QY 246 -DMVMFLVCVVTFITAFGAATAYGLYLVRNGSIDPDATAAPVPAGTTKAEAAEAPFS 304
Db 420 SNIVNALIVAVISVITFVLAIYFG-----IDEESS-----SNLEVGAGVS 461

QY 305 NDSITIQAPLTGEATLSVSDAMFASGKLGSGVAIVPTKGOLVSPVSGKIIVAPFSGHA 364
Db 462 NKKMIF-SPISGEIIPLSVDQKTFSDKLGDLGVALIPSEGVYAPFDGKITNIPTKHA 520

QY 365 FAVRKAEDGNSVDILMHIGFTVNLNGTHFPLKQGVKAGELLCEFFDIDAIAKAGY 424
Db 521 IGL--KSEGG--VELLIHGLTVELKGGFTSHVEEGRVFNQILFEMDLNLKTKGY 576

QY 425 EYVTTPIVNSN 434
Db 577 EVNTPVIVTN 586

RESULT 9
US-09-543-681A-7019
; Sequence 7019, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETTON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 7019
; LENGTH: 634
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-7019

Query Match      24.1%; Score 569; DB 4; Length 634;
Best Local Similarity 32.8%; Pred. No. 1.6e-48;
Matches 150; Conservative 90; Mismatches 161; Indels 56; Gaps 16;

QY 1 NAMVPPSLVNGYDVAATMAAGEMPWLSFLGDVAQAGYQGTGTVLPVLVSVWILATIEKFLH 60
Db 186 LAMVSP-----EWTSLVKAEG-PV-EPMHPVLVALVKSSQLIIPALITVMMWSIERFIV 237

QY 61 KRLKGTADFLITPVLTLLLTGTLFTIATGPAMRWGVDVLAHGLQGLYDFGPGVGLLFLGL 120
```

```
Db 238 RIVPENVKVMVPLLVILYSTPIALIAVGPTVSFAQLIADGVLLIQHTGTFIAIPLIVA 297
QY 121 VYSPVITGLHOSFPPIELFLNQGG--SFIFATASMANIAQGAACLAFFFLAKSEKLGK 178
Db 298 IYPWLVSIGNHKALSPVSIMLVEQKGFDPPIIRVMALCSNMQAAASLAVSVRTKNKTLKQ 357
QY 179 LAGASGVSAVL-GITEPAIFGNLRLRWPFFIGTGTAAIGGALIALF-----NIKAVAGLA 233
Db 358 LAFSASITAFFGITTEPAMYGVLKUKPMY-----ACMIGGAIAGLFAGIVKLFKAFYVVT 413
QY 234 AGEFLGV---VSIDAPMMVFLVCVVTFITAFGAATAYGLYLVRNGSIDPDATAAPV-- 288
Db 414 PGLSLPMMIS-DTDNQV---VNAITLILASVATFIATLII-----GFDDP--TDDPIRD 463
QY 289 -----PAGTTK-----AEAAEAPAEFSNDSTIIQAPLTGEATLSVSDAMFASGKLGSG 337
Db 464 EBENNKQAASNTKKPQIANSKLPVG-----LISPLQKTVALSEVNDTFASGIMPG 516
QY 338 VAIVPTKGOLVSPVSGKIIVVAPFSGHAFVRKKAEDGNSVDILMHIGFTVNLNGTHFNP 397
Db 517 MALIPTTGKVIAPADGVVDITFSSGHAIGTLV---NNIEMLIHVIGIDTVNLAGQHFTC 572
QY 398 LKKQGVDEKAGELLCEFFDIDAIAKAGYEVTTPIVNSN 434
Db 573 CVVKGQKVTGDTLAEFFDLDAIIAGYDPTMTIITN 609
```

```
RESULT 10
US-09-107-532A-3902
; Sequence 3902, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 3902:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 631 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
```


FAIR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342

366 VTEGQVTPDTLVAEADLEAIKAAGKETSMLVLTNNMDRVKNFVLEXTGAKAKAPNDV 425


```
; SEQ ID NO 13018
; LENGTH: 483
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-13018

Query Match      20.7%; Score 489; DB 4; Length 483;
Best Local Similarity 38.5%; Pred.No. 1.3e-40;
Matches 104; Conservative 51; Mismatches 101; Indels 14; Gaps 3;

QY      3 MVFSLVNGYDVAA--TMAAGEMPWWSLFGLDVAQAGYQGTVPVLPVLVWSWILATIEKFL 59
Db      213 LTHPALTNANGVAAGFHTM-----NFFGIEVAMIGYQGTVPFVLLAVWFMSWEKRL 264

QY      60 HKRLKGTADFLITPVLTLTLTGFTFTFAIGPAMRWGDVLAHGLQGLYDFGPGVGGLLFG 119
Db      265 RRVIPDALDLITLPFLTVIISGFTIALLLIGPAGRALGDGTSFILSTLIISHAGWLAGLLFG 324

QY      120 LVYSPVITGLHQSPFPPIELELFNQ--GGSFIEATASMANIAQGAACLAVFFLAKSEKL 176
Db      325 GLYSVIVITGLHHSFHAIEAGLLGNPSIGVNFILPIWAMNVAQGGACFAVWFKTKDAKI 384

QY      177 KGLAGAGSVSAVLGITPFAIFGVNLRLEWPPFICIGTAAIGCALIALFNKIKAVALGAAGF 236
Db      385 KAITLPSAFSAMLGITTEAAIFGINLRFVKPTIAALVGGAGGAWVWSMHVYNTAVGLTAI 444

QY      237 LGVVSIDAPDMWMLVCAVVTFFIAFGNAI 266
Db      445 PGMAIVQASSLLNYIIGNAIAFAVAFALSL 474
```

Search completed: October 29, 2004, 23:15:53
Job time : 41 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 29, 2004, 22:57:48 ; Search time 156 Seconds

(without alignments)
1076.188 Million cell updates/sec

Title: US-09-604-231-2

Perfect score: 2363

Sequence: 1 MAWVPSLVNGYDVATMAA.....IEAGANLLNVAKEAVPATP 468

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_23Sep04:*

1: Geneseqp1980s:*

2: Geneseqp1990s:*

3: Geneseqp2000s:*

4: Geneseqp2001s:*

5: Geneseqp2002s:*

6: Geneseqp2003as:*

7: Geneseqp2003bs:*

8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	2363	100.0	468	4	AAB66707
2	2363	100.0	468	5	ABG80325
3	2363	100.0	468	5	AAG92850
4	2332	98.7	661	4	AAB69080
5	1821	77.1	362	4	AAB66708
6	832.5	35.2	496	7	ADH87937
7	695	29.4	617	5	ABH47495
8	695	29.4	617	6	ABU32951
9	673	28.5	655	6	ABU02378
10	673	28.5	655	6	ABP81458
11	669	28.3	627	8	ADK46609
12	667	28.2	634	5	ABH49823
13	663.5	28.1	676	5	ABP25654
14	662	28.0	628	7	ADU95661
15	660	27.9	627	4	AU37874
16	659	27.9	627	6	ABM72990
17	655	27.7	655	8	ADK46443
18	647.5	27.4	674	5	ABP25655
19	632	26.7	620	5	ABP27216
20	628	26.6	617	5	ABH49923
21	626.5	26.5	656	7	ABO61695
22	609	25.8	670	7	ADH86721
23	591.5	25.0	639	5	ABP27215
24	588.5	24.9	636	5	ABH54803
25	585	24.8	381	6	ABP81450

26	583.5	24.7	583	7	ADC97184	E. faeciu
27	580	24.5	618	5	ABH47649	Listeria
28	579	24.5	622	5	ABP28755	Streptoco
29	578.5	24.5	612	6	ABU45919	Protein e
30	572.5	24.2	612	8	ADK46396	Streptoco
31	570.5	24.1	612	6	ABU00950	S. pneumo
32	570.5	24.1	612	6	ABP81336	Streptoco
33	569	24.1	634	7	ADP06734	Bacterial
34	564.5	23.9	620	5	ABP26854	Streptoco
35	564.5	23.9	620	6	ABU46515	Protein e
36	562.5	23.8	631	7	ADC94275	E. faeciu
37	538	22.8	618	8	ADN61668	Corynebac
38	537	22.7	316	7	ADH87597	Enterococ
39	530	22.4	334	6	ABU02198	S. pneumo
40	520.5	22.0	243	5	ABP39289	Staphyloc
41	509	21.5	640	5	ABH49939	Listeria
42	501.5	21.2	427	7	ADC95903	E. faeciu
43	497.5	21.1	526	7	ADH86830	Enterococ
44	492	20.8	379	6	ABU42285	Protein e
45	492	20.8	478	6	ABM71603	Staphyloc

ALIGNMENTS

RESULT 1

AAB66707
ID AAB66707 standard; protein; 468 AA.

XX
AC AAB66707;

XX
DT 09-APR-2001 (first entry)

XX
DE C.Glutamicum phosphoenolpyruvate protein #1.

XX
KW Phosphoenolpyruvate; sugar phosphotransferase system; PTS.

XX
OS Corynebacterium glutamicum.

XX
PN WO200102583-A2.

XX
PD 11-JAN-2001.

XX
PF 27-JUN-2000; 2000WO-IB000973.

XX
PR 01-JUL-1999; 99US-0142691P.

XX
PR 23-AUG-1999; 99US-0150310P.

XX
PR 03-SEP-1999; 99DE-01042095.

XX
PR 03-SEP-1999; 99DE-01042097.

XX
(BADI) BASF AG.

XX
Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;

XX
WPI; 2001-080989/09.

XX
Corynebacterium glutamicum nucleic acids encoding phosphoenolpyruvate:
sugar phosphotransferase system proteins or their portions, useful for
typing or identifying C. glutamicum or related bacteria, and as markers
for transformation.

XX
Claim 4; Page 101-102; 144pp; English.

XX
The present invention relates to Corynebacterium glutamicum
phosphoenolpyruvate: sugar phosphotransferase system (PTS) proteins. The
PTS nucleic acids and proteins are useful in the identification of
microorganisms which can be used to produce fine chemicals, for
modulating fine chemical production in C. glutamicum or related bacteria,
the typing or identification of C. glutamicum or related bacteria, as
reference points for mapping C. glutamicum genome, and as markers for
transformation

XX
SQ Sequence 468 AA;

Query Match		100.0%;	Score 2363;	DB 4;	Length 468;
Best Local Similarity		100.0%;	Pred. NO. 1.2e-226;		
Matches 468;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MAMVFP	SLVNGYDVAATMAAGEMPWMSLFGLDVAQAQYQGTGTLVPLVSVWILATIEKFLH	60	
DB	1	MAMVFP	SLVNGYDVAATMAAGEMPWMSLFGLDVAQAQYQGTGTLVPLVSVWILATIEKFLH	60	
QY	61	KRLKGT	ADFLITPVLITLLTGFTLFTFIAIGPAMRWGVDVLAHGLQGLYDFGPGVGLLFG	120	
DB	61	KRLKGT	ADFLITPVLITLLTGFTLFTFIAIGPAMRWGVDVLAHGLQGLYDFGPGVGLLFG	120	
QY	121	VYSPIV	TGLHQSPPPIELFNQGSFIFATASMANIAQGAACIAVFFLAKSEKLG	180	
DB	121	VYSPIV	TGLHQSPPPIELFNQGSFIFATASMANIAQGAACIAVFFLAKSEKLG	180	
QY	181	GASGVS	AVLGITTEPAIFGVNLRWPPFIFIGTAAIGGALIALFNKAVALLVAAGAGFLGV	240	
DB	181	GASGVS	AVLGITTEPAIFGVNLRWPPFIFIGTAAIGGALIALFNKAVALLVAAGAGFLGV	240	
QY	241	SIDAPD	WMFLVCAVWTFPIAFGAIAIYGLVLRNGSIDPDATAAPVPACTTKAEAP	300	
DB	241	SIDAPD	WMFLVCAVWTFPIAFGAIAIYGLVLRNGSIDPDATAAPVPACTTKAEAP	300	
QY	301	AEFSND	STIIQAPLTGEAIALSSVSDAMFASGKLGSGVAIVPTKGOLVSPVSGKI	360	
DB	301	AEFSND	STIIQAPLTGEAIALSSVSDAMFASGKLGSGVAIVPTKGOLVSPVSGKI	360	
QY	361	SGHAFV	TKAEDGNSVDILMHIGDVTNLTGTHFNPLKQGDVKAAGELLCEFDIDA	420	
DB	361	SGHAFV	TKAEDGNSVDILMHIGDVTNLTGTHFNPLKQGDVKAAGELLCEFDIDA	420	
QY	421	AAGYEV	TPPIVVSNNKTKGPVNTYGLGEIEAGANLLINVAKEAVPATP	468	
DB	421	AAGYEV	TPPIVVSNNKTKGPVNTYGLGEIEAGANLLINVAKEAVPATP	468	
RESULT 2					
ABG80325					
ID	ABG80325 standard; protein; 468 AA.				
XX	AC ABG80325;				
XX	DT 15-NOV-2002 (first entry)				
XX	C. glutamicum metabolic pathway (MP) protein #5.				
XX	Metabolic pathway protein; MP; cell metabolism; amino acid; vitamin;				
KW	cofactor; nucleotide; nucleoside; trehalose; fine chemical production;				
KW	organic acid; non-proteinogenic amino acid; purine base; carboxylate;				
KW	pyrimidine base; lipid; unsaturated fatty acid; diol; polyketide;				
KW	aromatic compound; food industry; animal feed; cosmetic industry;				
KW	pharmaceutical industry; enzyme.				
XX	Corynebacterium glutamicum ATCC 13032.				
OS	WO200251231-A1.				
XX	04-JUL-2002.				
XX	22-DEC-2000; 2000WO-EP013143.				
XX	22-DEC-2000; 2000WO-EP013143.				
XX	(BADI) BASF AG.				
XX	Pompejus M, Kroeger B, Zelder O, Schroeder H;				
PI	WPI; 2002-643289/69.				
DR	N-PSDB; ABS65346.				
XX	New metabolic pathway genes of Corynebacterium glutamicum for producing				
PT					
PT	fine chemicals, e.g. lipids, (un)saturated fatty acids, vitamins,				
PT	cofactors or enzymes used in food, feed, cosmetics or pharmaceutical				
PT	industries.				
XX	Claim 18; Page 111-113; 176pp; English.				
XX	The present invention relates to the isolation of Corynebacterium				
CC	glutamic metabolic pathway (MP) proteins, and the polynucleotide				
CC	sequences encoding them. The MP proteins are enzymes involved in the				
CC	metabolism of molecules important for the normal functioning of cells				
CC	(e.g. amino acids, vitamins, cofactors, nucleotides and nucleosides, or				
CC	trehalose). The polynucleotide sequences encoding the MP proteins are				
CC	useful for producing fine chemicals, particularly organic acids, non-				
CC	proteinogenic amino acids, purine and pyrimidine bases, nucleosides,				
CC	nucleotides, lipids, (un)saturated fatty acids, diols, carboxylates,				
CC	aromatic compounds, vitamins, cofactors, polyketides and enzymes. The				
CC	fine chemicals are useful in the food, animal feed, cosmetic or				
CC	pharmaceutical industries. ABG80321-ABG80343 represent the C. glutamicum				
CC	MP proteins of the invention				
XX	Sequence 468 AA;				
Query Match		100.0%;	Score 2363;	DB 5;	Length 468;
Best Local Similarity		100.0%;	Pred. NO. 1.2e-226;		
Matches 468;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MAMVFP	SLVNGYDVAATMAAGEMPWMSLFGLDVAQAQYQGTGTLVPLVSVWILATIEKFLH	60	
DB	1	MAMVFP	SLVNGYDVAATMAAGEMPWMSLFGLDVAQAQYQGTGTLVPLVSVWILATIEKFLH	60	
QY	61	KRLKGT	ADFLITPVLITLLTGFTLFTFIAIGPAMRWGVDVLAHGLQGLYDFGPGVGLLFG	120	
DB	61	KRLKGT	ADFLITPVLITLLTGFTLFTFIAIGPAMRWGVDVLAHGLQGLYDFGPGVGLLFG	120	
QY	121	VYSPIV	TGLHQSPPPIELFNQGSFIFATASMANIAQGAACIAVFFLAKSEKLG	180	
DB	121	VYSPIV	TGLHQSPPPIELFNQGSFIFATASMANIAQGAACIAVFFLAKSEKLG	180	
QY	181	GASGVS	AVLGITTEPAIFGVNLRWPPFIFIGTAAIGGALIALFNKAVALLVAAGAGFLGV	240	
DB	181	GASGVS	AVLGITTEPAIFGVNLRWPPFIFIGTAAIGGALIALFNKAVALLVAAGAGFLGV	240	
QY	241	SIDAPD	WMFLVCAVWTFPIAFGAIAIYGLVLRNGSIDPDATAAPVPACTTKAEAP	300	
DB	241	SIDAPD	WMFLVCAVWTFPIAFGAIAIYGLVLRNGSIDPDATAAPVPACTTKAEAP	300	
QY	301	AEFSND	STIIQAPLTGEAIALSSVSDAMFASGKLGSGVAIVPTKGOLVSPVSGKI	360	
DB	301	AEFSND	STIIQAPLTGEAIALSSVSDAMFASGKLGSGVAIVPTKGOLVSPVSGKI	360	
QY	361	SGHAFV	TKAEDGNSVDILMHIGDVTNLTGTHFNPLKQGDVKAAGELLCEFDIDA	420	
DB	361	SGHAFV	TKAEDGNSVDILMHIGDVTNLTGTHFNPLKQGDVKAAGELLCEFDIDA	420	
QY	421	AAGYEV	TPPIVVSNNKTKGPVNTYGLGEIEAGANLLINVAKEAVPATP	468	
DB	421	AAGYEV	TPPIVVSNNKTKGPVNTYGLGEIEAGANLLINVAKEAVPATP	468	
RESULT 3					
AAG92650					
ID	AAG92650 standard; protein; 661 AA.				
XX	AC AAG92650;				
XX	DT 26-SEP-2001 (first entry)				
XX	C glutamicum protein fragment SEQ ID NO: 6404.				
XX	Coryneform bacterium; amino acid synthesis; vitamin; saccharide;				
KW	organic acid synthesis.				
XX	Corynebacterium glutamicum.				

XX PN EP1108790-A2.
 XX PD 20-JUN-2001.
 XX PF 18-DEC-2000; 2000EP-00127688.
 XX PR 16-DEC-1999; 99JP-00377484.
 XX PR 07-APR-2000; 2000JP-00159162.
 XX PR 03-AUG-2000; 2000JP-00280988.
 XX PA (KYOW) KYOWA HAKKO KOGYO KK.
 XX PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
 XX PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
 XX DR WPI; 2001-376931/40.
 XX DR N-PSDB; AAH67869.
 XX PT Novel polynucleotides derived from Coryneform bacteria, for identifying
 XX PT mutation point of a gene, measuring expression of a gene, analyzing
 XX PT expression profile or pattern of a gene and identifying homologous gene.
 XX PS Claim 17; SEQ ID NO 6404; 246pp + Sequence Listing; English.
 XX CC The present invention provides a number of nucleotide and protein
 XX CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
 XX CC are useful for identifying the mutation point of a gene derived from a
 XX CC mutant of coryneform bacterium, measuring expression amount and analyzing
 XX CC the expression profile or expression pattern of a gene derived from
 XX CC Coryneform bacterium, and identifying a homologue of a gene derived from
 XX CC coryneform bacterium. Coryneform bacteria are useful for producing amino
 XX CC acids, nucleic acids, vitamins, saccharides and organic acids,
 XX CC particularly L-lysine. The present sequence is a protein described in the
 XX CC exemplification of the invention. Note: The sequence data for this patent
 XX CC did not form part of the printed specification, but was obtained in
 XX CC electronic format directly from the European Patent Office
 XX SQ Sequence 661 AA;
 Query Match 100.0%; Score 2363; DB 4; Length 661;
 Best Local Similarity 100.0%; Pred. No. 2.1e-226;
 Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MAMVPPSLVNGYDVAAATMAAGEMPWMSLFGLDVAQAGYQGTIVLVVSVWILATIEKFLH 60
 DB 194 MAMVPPSLVNGYDVAAATMAAGEMPWMSLFGLDVAQAGYQGTIVLVVSVWILATIEKFLH 253
 QY 61 KRLKGTADFLITPVLTLLTGLTFPIAIGPAMRWVGDVLAHGLQGLYDFGPGVGLLFL 120
 DB 254 KRLKGTADFLITPVLTLLTGLTFPIAIGPAMRWVGDVLAHGLQGLYDFGPGVGLLFL 313
 QY 121 VYSPITVITGLHQSPPPIELELFNQGSGSFIATASMANIAQGAACLAFFLAKSEKLGKLA 180
 DB 314 VYSPITVITGLHQSPPPIELELFNQGSGSFIATASMANIAQGAACLAFFLAKSEKLGKLA 373
 QY 181 GASGVSALVIGTEPAIFGNVLRWPPFPGICTAIGGALLIALFNIKAVALCAAGFLGV 240
 DB 374 GASGVSALVIGTEPAIFGNVLRWPPFPGICTAIGGALLIALFNIKAVALCAAGFLGV 433
 QY 241 SIDAPDMVFLVCAVTFPIAAGAAIAYGLVLRNNGSIDPDAAAPVPGTTKAEAEAP 300
 DB 434 SIDAPDMVFLVCAVTFPIAAGAAIAYGLVLRNNGSIDPDAAAPVPGTTKAEAEAP 493
 QY 301 AEFNSDSTIIIOAPLTGEAIALSSVSDAMFASGKLGSGVAIVPTKGOLVSPVSGKIVVAP 360
 DB 494 AEFNSDSTIIIOAPLTGEAIALSSVSDAMFASGKLGSGVAIVPTKGOLVSPVSGKIVVAP 553
 QY 361 SGHAFVTRKAEAGSNVDILMHI GFDTVNLTNGTHFNPLKKGDEKAGELLCEFDIDA 420
 DB 554 SGHAFVTRKAEAGSNVDILMHI GFDTVNLTNGTHFNPLKKGDEKAGELLCEFDIDA 613
 QY 421 AAGYEVTPIVVSNNKKTGPVNTYGLGEIAGANLLNVAKKEAVPATP 468

DB 614 AAGYEVTPIVVSNNKKTGPVNTYGLGEIAGANLLNVAKKEAVPATP 661
 RESULT 4
 AAB69080
 ID AAB69080 standard; protein; 661 AA.
 XX AC AAB69080;
 XX DT 11-SEP-2003 (revised)
 XX DT 20-APR-2001 (first entry)
 XX Brevibacterium lactofermentum sucrose PTS enzyme II protein SEQ ID NO:2.
 XX Brevibacterium lactofermentum; sucrose PTS enzyme II; sucrose;
 XX Brevibacterium lactofermentum; sucrose PTS enzyme II; sucrose;
 XX coryneform bacterium; carboxylate phosphotransferase system; glucose;
 XX coryneform bacterium; phosphoenolpyruvate-sugar transport system.
 XX OS Corynebacterium glutamicum.
 XX PN WO200102584-A1.
 XX PD 11-JAN-2001.
 XX PF 30-JUN-2000; 2000WO-JP004348.
 XX PR 02-JUL-1999; 99JP-00189512.
 XX PA (AJIN) AJINOMOTO CO-INC.
 XX PI Izui M, Sugimoto M, Nakamatsu T, Kurahashi O;
 XX WPI; 2001-138150/14.
 XX DR N-PSDB; AAF32543.
 XX PT Gene encoding sucrose phosphoenolpyruvate-sugar transport system enzyme
 XX PT II obtained by cassette ligation-mediated amplification of downstream
 XX PT domain of coryneform bacterium sucrose gene, with sucrose-binding
 XX PT activity.
 XX PS Claim 1; Page 29-32; 45pp; Japanese.
 XX CC The present sequence represents the Brevibacterium lactofermentum sucrose
 XX CC PTS (phosphoenolpyruvate:carboxylate phosphotransferase system or
 XX CC phosphoenolpyruvate-sugar transport system) enzyme II, which has sucrose-
 XX CC binding activity. A coryneform bacteria produced with the sucrose PTS
 XX CC enzyme II gene can have more efficient sugar uptake, and improved amino-
 XX CC acid and nucleic acid productivity. The sucrose PTS gene and its
 XX CC disrupted gene, such as one without the sucrose PTS function, can be used
 XX CC to produce new breeds of coryneform bacterial strains to uptake sugar
 XX CC more efficiently e.g. glucose only or and sucrose, and can have improved
 XX CC amino-acid and nucleic acid productivity. (Updated on 11-SEP-2003 to
 XX CC standardise OS field)
 XX SQ Sequence 661 AA;
 Query Match 98.7%; Score 2332; DB 4; Length 661;
 Best Local Similarity 98.5%; Pred. No. 2.6e-223;
 Matches 461; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
 QY 1 MAMVPPSLVNGYDVAAATMAAGEMPWMSLFGLDVAQAGYQGTIVLVVSVWILATIEKFLH 60
 DB 194 MAMVPPSLVNGYDVAAATMAAGEMPWMSLFGLDVAQAGYQGTIVLVVSVWILATIEKFLH 253
 QY 61 KRLKGTADFLITPVLTLLTGLTFPIAIGPAMRWVGDVLAHGLQGLYDFGPGVGLLFL 120
 DB 254 KRLKGTADFLITPVLTLLTGLTFPIAIGPAMRWVGDVLAHGLQGLYDFGPGVGLLFL 313
 QY 121 VYSPITVITGLHQSPPPIELELFNQGSGSFIATASMANIAQGAACLAFFLAKSEKLGKLA 180
 DB 314 VYSPITVITGLHQSPPPIELELFNQGSGSFIATASMANIAQGAACLAFFLAKSEKLGKLA 373

```
QY 181 GASGVAVLIGITPAIFGVNLRWRPFYIGTAAIGGALLIALFNLIKAVAGAGFLGVV 240
DB 374 GASGVAVLIGITPAIFGVNLRWRPFYIGTAAIGGALLIALFDIKAVAGAGFLGVV 433
QY 241 SIDAPDMVFLVCAVVTFFIAFGAAIAYGLYLVRNGSIDPDATAAPVPAGTTKAEAEAP 300
DB 434 SIDAPDMVFLVCAVVTFFIAFGAAIAYGLYLVRNGSIDPDATAAPVPAGTTKAEAEAP 493
QY 301 AEFNSDSTIIQAPLTGEAIALSSVSDAMFASGKLGSGVAIVPTKGQIVPSVSGKIIVVAFP 360
DB 494 AEFNSDSTIIQAPLTGEAIALSSVSDAMFASGKLGSGVAIVPTKGQIVPSVSGKIIVVAFP 553
QY 361 SCHAFVRTKAEDGNSVDILMHIGFDVNLNGTHFNPLKKGDSVKAGELLCEFDIDAIAK 420
DB 554 SCHAFVRTKAEDGNSVDILMHIGFDVNLNGTHFNPLKKGDSVKAGELLCEFDIDAIAK 613
QY 421 AAGVEVTTPIVWSNYKKTGPVNTYGLGEIEAGANLLNVAKKEAVPATP 468
DB 614 AAGVEVTTPIVWSNYKKTGPVNTYGLGEIEAGANLLNVAKKEAVPATP 661

RESULT 5
AAB66708
ID AAB66708 standard; protein; 362 AA.
AC AAB66708;
XX
XX
XX 09-APR-2001 (first entry)
XX
XX C.glutamicum phosphoenolpyruvate protein #2.
XX
XX Phosphoenolpyruvate; sugar phosphotransferase system; PTS.
XX
XX Corynebacterium glutamicum.
XX
XX WO200102583-A2.
XX
XX 11-JAN-2001.
XX
XX 27-JUN-2000; 2000WO-IB000973.
XX
XX 01-JUL-1999; 99US-0142691P.
XX
XX 23-AUG-1999; 99US-0150310P.
XX
XX 03-SEP-1999; 99DE-01042095.
XX
XX 03-SEP-1999; 99DE-01042097.
XX
XX (BADI ) BASF AG.
XX
XX Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;
XX
XX WPI; 2001-080989/09.
XX
XX Corynebacterium glutamicum nucleic acids encoding phosphoenolpyruvate:
XX sugar phosphotransferase system proteins or their portions, useful for
XX typing or identifying C. glutamicum or related bacteria, and as markers
XX for transformation.
XX
XX Claim 6; Page 104-106; 144pp; English.
XX
XX The present invention relates to Corynebacterium glutamicum
XX phosphoenolpyruvate; sugar phosphotransferase system (PTS) proteins. The
XX PTS nucleic acids and proteins are useful in the identification of
XX microorganisms which can be used to produce fine chemicals, for
XX modulating fine chemical production in C. glutamicum or related bacteria,
XX the typing or identification of C. glutamicum or related bacteria, as
XX reference points for mapping C. glutamicum genome, and as markers for
XX transformation
XX
XX Sequence 362 AA;
XX
XX Query Match 77.1%; Score 1821; DB 4; Length 362;
XX Best Local Similarity 100.0%; Pred. No. 1.1e-172;
XX Matches 362; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```


PT ear infection.

PS Claim 1; SEQ ID NO 3912; 56pp; English.

XX The invention relates to a protein comprising or having at least 50%
 CC identity to any of the 2469 amino acid sequences, identified in the
 CC specification (available on a computer readable format), or its fragment,
 CC expressed from 2469 of 2489 identified DNA coding regions from the
 CC Streptococcus pneumoniae type 4 strain genomic sequence appearing as
 CC ABS56454. Also included are an antibody which binds one of the proteins,
 CC treating a patient by administering the protein, DNA or antibody (in a
 CC composition), a kit comprising first and second primers, which are the
 CC nucleic acid cited above or fragments between nucleotides 8-100 of a
 CC sequence not defined in the specification, for amplifying a target
 CC the first primer is substantially complementary to the target sequence
 CC and the second primer is substantially complementary to the target sequence
 CC the target sequence, and where the parts of the primers having
 CC substantial complementarity define the termini of the target sequence to
 CC be amplified, assay comprising contacting a test compound with the
 CC protein, and determining whether the test compound binds to the protein
 CC and a Streptococcus pneumoniae bacterium, where one or more genes
 CC encoding the proteins has been rendered inactive. The proteins, nucleic
 CC acid molecules, antibody and compositions are useful as medicaments for
 CC treating or preventing a disease or infection due to Streptococcus
 CC bacteria, particularly S. pneumoniae, such as pneumonia, sepsis, otitis
 CC media or ear infection. They are also useful in developing vaccines,
 CC diagnostics and antibiotics. The methods are useful for identifying
 CC immunodominant proteins. The present sequence is one of the 2469 proteins
 CC expressed by the identified coding regions from the genomic sequence.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences. (Updated on 23-OCT-2003 to
 CC standardise OS field)

XX Sequence 655 AA;

Query Match 28.5%; Score 673; DB 6; Length 655;
 Best Local Similarity 35.3%; Pred. No. 1.1e-57;
 Matches 170; Conservative 87; Mismatches 173; Indels 59; Gaps 16;
 QY 1 MAMVFPVLVNGYDVAATWAAGEMP--MWSFLGLDVAQAGYQGVLPVLVSWILATIEKF 58
 DB 195 ICLVSPQLLNAYAVASTPAADIAANWVWNGFYFTVNRIGYQAQVIPALLAGLSLYLEIF 254
 QY 59 LHKRLKGTADFLITPVLTLTLLTGTFTFIAIGPMRWGVDVLAHQGLGYDFG--GPVG-- 114
 DB 255 WKHHPVISMIFVFPFLSLIPALLAHTVLGP-----IGWTIGQGLSSVVLAGLTGPVKWL 310
 QY 115 -GLLFLGLVSPVITGLHQSPPPIELELF--NQGSFIFATASMANIAQGAACLAFFFLAK 172
 DB 311 FGAIFGALYAPFVITGLHHTNAIDTQLIADAGGTALWPMIALSNIAQGSVAFFAYFMR 370
 QY 173 -SEKLKGLAGASGVSVLGTITEPAIFGVNLRWLPFPFFIGTAAIGGALIALFNKVAL 231
 DB 371 HDREAQVSLPATISALIGVTEPALFGVNKYIYPFVAGMTGSGALAGMLSVTFNVTRAISI 430
 QY 232 GAAGFLGVVSDIDPDMVFP-----LVCAVVTFFIAFGAAIAYGLVLRNGSIDPDATAAP 287
 DB 431 GIGGLPGLISLPQYMLPFAGTMLVAIVVPMLLTF-----FFRAGLFTK----- 475
 QY 288 VPAGTTAAEAE-----APAEFSNDGTI-----IOAPLTGEAIALSSVSDAMFASCKLGS 337
 DB 476 -TEGDTMLQAEFAQEAEEFVNHPEVELTSVEIISPLTGQVKELSQATDPIFASGVWGQG 534
 QY 338 VAIPTKGLQVSPVSGKIVVAPFSGHAFATKEDGSDVDILMHIGDVTNLNGTHFNP 397
 DB 535 LVIEPSGELTSPVNGVTVLFPFKTHAIGI--VSDEG--VELLIHIGMDTVGLDGKGFES 590
 QY 398 LKQGDVVKAGELLCEPDIIDAIKAAGYEVTTPIVVSN---YKKTGPVNTYGLGHIKAGAN 454
 DB 591 LVVQGDHVTVGOQLIRFDMVDVKAAGLVTEPVTPIITNQDAYTATIP--GTPT--TIQAGAS 648

QY 455 LL 456
 DB 649 LM 650

RESULT 10

ABP81458
 ID ABP81458 standard; protein; 655 AA.

XX AC ABP81458;

XX DT 04-MAR-2003 (first entry)

XX Streptococcus pneumoniae polypeptide SEQ ID NO 375.

XX Streptococcus pneumoniae; infection; otitis media; antibacterial;
 KW diagnosis; gene therapy.

XX Streptococcus pneumoniae.

XX WO200283855-A2.

XX 24-OCT-2002.

XX 12-APR-2002; 2002WO-US011524.

XX 16-APR-2001; 2001US-0283948P.

XX 18-APR-2001; 2001US-0284443P.

XX (AMCY) AMERICAN CYANAMID CO.

XX Zagursky RJ, Masi AW, Green BA, Chakravarti DN, Russell DP;

PI Wothers JL;

XX WPI; 2003-093010/08.

XX N-PSDB; AB242306.

XX New Streptococcus pneumoniae polynucleotides, useful for treating or
 PT preventing S. pneumoniae infections, or non-systemic diseases, e.g.
 PT otitis media, which are induced or exacerbated by S. pneumoniae.

XX Claim 42; Page 617-620; 1091pp; English.

XX The invention relates to isolated polynucleotides (AB272147-AB242522) of
 CC a Streptococcus pneumoniae genomic sequence, a fragment or degenerate
 CC variant of the polynucleotide or a nucleic acid sequence 95% identical to
 CC one of the polynucleotides. The S. pneumoniae polynucleotides and encoded
 CC polypeptides (ABP81299-ABP81674) are useful for treating or preventing S.
 CC pneumoniae infections or non-systemic diseases, e.g. otitis media, which
 CC are induced or exacerbated by S. pneumoniae. These are also useful for
 CC detecting S. pneumoniae in a biological sample or diagnosing S.
 CC pneumoniae infection in a subject. The polynucleotides have antibacterial
 CC activity and are useful in gene therapy

XX Sequence 655 AA;

Query Match 28.5%; Score 673; DB 6; Length 655;
 Best Local Similarity 35.3%; Pred. No. 1.1e-57;
 Matches 170; Conservative 87; Mismatches 173; Indels 52; Gaps 16;
 QY 1 MAMVFPVLVNGYDVAATWAAGEMP--MWSFLGLDVAQAGYQGVLPVLVSWILATIEKF 58
 DB 195 ICLVSPQLLNAYAVASTPAADIAANWVWNGFYFTVNRIGYQAQVIPALLAGLSLYLEIF 254
 QY 59 LHKRLKGTADFLITPVLTLTLLTGTFTFIAIGPMRWGVDVLAHQGLGYDFG--GPVG-- 114
 DB 255 WKHHPVISMIFVFPFLSLIPALLAHTVLGP-----IGWTIGQGLSSVVLAGLTGPVKWL 310
 QY 115 -GLLFLGLVSPVITGLHQSPPPIELELF--NQGSFIFATASMANIAQGAACLAFFFLAK 172
 DB 311 FGAIFGALYAPFVITGLHHTNAIDTQLIADAGGTALWPMIALSNIAQGSVAFFAYFMR 370
 QY 173 -SEKLKGLAGASGVSVLGTITEPAIFGVNLRWLPFPFFIGTAAIGGALIALFNKVAL 231

Db 371 HDEREQVSLPATISAYLGVTEPAGLVNKKYIYPPVAGMTGSLAGMLSVTENVTAASI 430
QY 232 GAAGFLGVSVISIDAPDMVF-----LVCVVTFFAFGAIAIYGLYLVRRNGSIDPDATAAP 287
Db 431 GIGGLPGLLSIQPYMLPFAGTMLVAIVPMLLTF-----FFRKAGLFTK----- 475
QY 288 VPAGTTKAEAE-----APAEHSNDSTI-----TQAPLTGEAIALSSVSDAMFASGKLGG 337
Db 476 -TEGDTNLOAEFAVQAEAEFVNHEPVELTSVEIISPLTGQVKELSQAIDPFIASGVWGQ 534
QY 338 VATPRTKQGLVSVSGKIVVAFPSGHAFVARTKAEDGNSVDIIMHIGDFTVNLNGTHFNP 397
Db 535 LVIEPSOGELTSPVNGTVTLFPTKHAIGI--VSDEG--VELLIHIGMDVGLDGKGFES 590
QY 398 LKQGGDEVKAGELCEPDIIDAIKAAGYEVVTPVVSNN---YKKTGPVNTYGLGEIEAGAN 454
Db 591 LVVQGDHVTVGQQLIRFDMVDVKAAGLVVTEPTVITNQDAYTATIP-GTYPT-TIQAGAS 648
QY 455 LL 456
Db 649 LM 650
RESULT 11
ADK46609
ID ADK46609 standard; protein; 627 AA.
XX
AC ADK46609;
XX
DT 20-MAY-2004 (first entry)
XX
DE Streptococcus pneumoniae protein, Seq ID No 3124.
XX
KW Antibacterial; Gene therapy; Vaccine; Streptococcus pneumoniae.
XX
OS Streptococcus pneumoniae.
XX
FN US6699703-B1.
XX
PD 02-MAR-2004.
XX
PF 26-MAY-2000; 2000US-00583110.
XX
PR 02-JUL-1997; 97US-0051553P.
PR 12-MAY-1998; 98US-0085131P.
PR 30-JUN-1998; 98US-00107433.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
XX
XX Doucette-Stamm L, Bush D, Zeng Q, Opperman T, Houseweart CE;
XX WPI; 2004-212399/20.
XX N-PSDB; ADK43948.
XX
XX New nucleic acid molecules and polypeptides useful for diagnosing,
PT preventing and treating pathological conditions resulting from bacterial
PT infection, e.g. Streptococcus pneumoniae infection, and in drug
PT screening.
XX
XX Disclosure; SEQ ID NO 3124; 301pp; English.
XX
XX The invention relates to isolated Streptococcus pneumoniae nucleic acids
CC and polypeptides. The nucleic acids and proteins are useful for
CC diagnosing, preventing and treating pathological conditions resulting
CC from bacterial infection, such as S. pneumoniae infection. These may also
CC be used for drug screening procedures. The present sequence represents a
CC Streptococcus pneumoniae polypeptide of the invention. Note: The sequence
CC data for this patent did not appear in the printed specification but was
CC obtained in electronic format directly from USPTO at
XX seqdata.uspto.gov/sequence.html.
XX
XX Sequence 627 AA;
SQ

Query Match 28.3%; Score 669; DB 8; Length 627;
Best Local Similarity 34.5%; Pred. No. 2,7e-57;
Matches 160; Conservative 92; Mismatches 186; Indels 26; Gaps 10;
QY 1 MANVPSLVNGYDVAATMAAGEMPWSLFLGLDVAQAGYQGTLPVLVWSWILATIKFLH 60
Db 182 MMLVSGSLPNMAVA---QGGEVTAMNFGF-IPVVLQGSVLPAFTIIVGVGAKFEKAVR 237
QY 61 KRUKGTADFLITPVLITLLTGTFTFAIGFAMRWGVDVLAHGLQGLYDFGCGPVGGLLFLG 120
Db 238 KVVDPDVIDLLVTFVILLVMSILGLFVGPVHVVENYILIAATKAILSIPLGLGGFTLGG 297
QY 121 VYSPVITGLHQSPFPIELELFPNOGGSFIP-ATASHANIAQGAACLAFFFLAKSEKLG 179
Db 298 VHQLIVVSGVHHIFNLLEVLQLAADHANPENAIITAAMTAQGAATVAVGVKTNPKUTL 357
QY 180 AGASGSAVLGITTEPAIFGVNLRIRWPFPTIGTAAIGGALIALFNIAKAVAGAAAGFLG- 238
Db 358 APPAALSAGLITTEPAIFGVNLRIRWPFPTIGTAAIGGALIALFNIAKAVAGAAAGFLG- 417
QY 239 VVSIDAPDMVFLVCAVVTFFIAFGAAIAYGLVLRNNGSIDPDATAAPVPAGTTKAEAE 298
Db 418 MLYVNGQLPQYLLMWAVSFALGFALTYMEGY-----EDEVDTAAAKQAEVAEKEE 470
QY 299 -APAEFSNDSTIQAPLTGEAIALSSVSDAMFASGKLGSVAIVPTKQLVSPVSGKIV 357
Db 471 VAPAAQLNETLIV--TPIVGDVVVALADVNDPVFSSGAMGQGIHVKPSQVVTAPADAESVI 528
QY 358 APPSGHAFVARTKAEDGNSVDIIMHIGDFTVNLNGTHFNPCLKQGDDEVKAGELCEFDID 417
Db 529 APFTGHAFGLKTR-----NGAEVLHVGLDITVSMNGDGFKAQVQGNKVGADVLGTFDN 584
QY 418 AKAAGYEVVTPVVSNN---YKKTGPVNTYGLGEIEAGANLLNV 458
Db 585 KIAAAGLDDTTMIVTNTADVASVAPVAT---GSVAKGDAVIEV 625
RESULT 12
ABB49833
ID ABB49833 standard; protein; 634 AA.
XX
AC ABB49833;
XX
DT 05-FEB-2002 (first entry)
XX
DE Listeria monocytogenes protein #2537.
XX
KW Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
KW vitamin B12; bacterial infection; disease.
XX
OS Listeria monocytogenes.
XX
PN WO200177335-A2.
XX
PD 18-OCT-2001.
XX
PF 11-APR-2001; 2001WO-FR001118.
XX
PR 11-APR-2000; 2000FR-00004629.
XX
XX (INSP) INST PASTEUR.
XX
XX Bucharier C, Frangeul L, Couve E, Rusniok C, Fsihi H, Dehoux P;
PI Dussurget O, Chetouani F, Nedjari H, Glaser P, Kunst F, Cossart P;
PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;
PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;
PI Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L;
PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;
PI Maduenio E, De Pablo B, Wehland J, Kaerst U, Entian K, Hauf J;
PI Rose M, Voss H;
XX
XX WPI; 2002-010914/01.
DR

XX Genomic sequence for *Listeria monocytogenes*, useful e.g. for treatment
 PT and prevention of *Listeria* and related bacterial infections, and related
 PT polypeptides.
 XX
 PS Claim 6; SEQ ID NO 2538; 192pp; French.
 XX
 CC The present invention relates to the genome sequence of *Listeria*
 CC *monocytogenes* EGD-e (see ABA03041). The genome sequence and fragments of
 CC it are useful for selecting probes and primers for detecting genes in *L.*
 CC *monocytogenes* and related organisms, and for studying genetic
 CC polymorphisms and other genomes. The present sequence is a protein
 CC encoded by the genome sequence of the present invention. Proteins
 CC expressed from the genome sequence are useful for raising specific
 CC antibodies, identification of *L. monocytogenes* and related organisms, and
 CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin
 CC B12. The genome sequence and proteins encoded by it are also useful for
 CC selecting compounds that regulate gene expression and cell replication
 CC and modulate *L. monocytogenes*-related diseases. In addition, the genome
 CC sequence and proteins encoded by it are useful in pharmaceutical and
 CC vaccine compositions for the treatment or prevention of infections by *L.*
 CC *monocytogenes* and related organisms. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 634 AA;

Query Match 28.2%; Score 667; DB 5; Length 634;
 Best Local Similarity 34.6%; Pred. No. 4,3e-57;
 Matches 156; Conservative 81; Mismatches 172; Indels 42; Gaps 10;
 QY 2 AMVFPSEL--VNGYDVAATMAAG---EMPWMSLF-GLDVAQAGYQGTLPVLVSVMLATPI 55
 DB 182 ALVPTLAGITAGDPLTLTFACTIFESPIHVTLGLPVLMSYASVPIIIVATVFGSKV 241
 QY 56 EKFLHRLKGTADFLLTPVLLTLLTGFTFTTIAIGPAMRWGVLDLAHGLQGLYDFGPGVG 115
 DB 242 EKGFKXIIIPDVIKTFVNPFCITLIVVPIITFVIGPIATWAGQLGAGTIWVNLSPITAG 301
 QY 116 LFLGLVSPVITGLHQSPPEL-ELNQGSGFTFATASMANIAQGAACLVFLAKSE 174
 DB 302 LILGFWQVVFVGLHWGLVPAVNNLTVLGHDPILAMTFGASFAQIGAVLAVFCKSRNK 361
 QY 175 KLKGLAGASGVSAVLGITTEPAIFGVNLRWRPFFIGIGTAAIGGALIALFALNIKAVALGAA 234
 DB 362 KIKSLISIPAFISGIPGVTPEALYGTLPKXPFINSIAGGILGFGASQTYIMGL 421
 QY 235 GFLGVVSDAP-----DMVPLVCVTFPIAFGAATAYGL-----VLRRNGSIDPDA 283
 DB 422 GIFGLNFFKPGSGISGFWVWVAIVISFTILGFTLTYVVGFKDPADVVEQSNVTE--- 478
 QY 284 TAAPVPAGTTKAEAEAPAEFSDNSTIIQAPITGAEIALSSVSDAMFASGKLGSGVAIVPT 343
 DB 479 -----GETLIERET-----TPAVVGVHVLADVKDEAFSGALGKGVAIPT 521
 QY 344 KQQLVSPVSGKIVVAFPSGHAFVTRKAEDGNSVDILMHIGFTVNLGTHFNPLKKGQD 403
 DB 522 VGRVVAAGAVTWTIFPTGHAIGITTK--DGA--EVLHIGNDIVQLGEGKFTAHVKGQD 577
 QY 404 EVKAGEILCEFDIDAIKAAGVETVPIVSN 434
 DB 578 VIERGQLITEFDIGIKAAAGYDVTTPVVVN 608

RESULT 13

ABP25654

ID ABP25654 standard; protein; 676 AA.

XX

AC ABP25654;

XX

DT 02-JUL-2002 (first entry)

XX

DE Streptococcus polypeptide SEQ ID NO 484.
 XX
 KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
 KW group A streptococcus; Streptococcus pyogenes; antibacterial;
 XX antiinflammatory; infection; vaccine; meningitis; gene therapy.
 OS Streptococcus agalactiae.
 PN WO200234771-A2.
 XX
 PD 02-MAY-2002.
 XX
 PF 29-OCT-2001; 2001WO-GB004789.
 XX
 PR 27-OCT-2000; 2000GB-00026333.
 PR 24-NOV-2000; 2000GB-00028727.
 PR 07-MAR-2001; 2001GB-00005640.
 XX
 PA (CHIR-) CHIRON SPA.
 PA (GENO-) INST GENOMIC RES.
 PI Telford J, Masignani V, Margarit Y Rosi, Grandi G, Fraser C;
 PI Tettelin H;
 XX
 DR WPI; 2002-352536/38.
 DR N-PSDB; ABN66285.
 XX
 XX New Streptococcus protein for the treatment or prevention of infection or
 PT disease caused by Streptococcus bacteria, such as meningitis, and for
 PT detecting a compound that binds to the protein.
 XX
 PS Claim 1; Page 3202-3203; 4525pp; English.
 XX
 CC The invention relates to a protein (ABP25413-ABP30895) from group B
 CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
 CC the specification. The proteins have antibacterial and antiinflammatory
 CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
 CC antibodies that bind (I) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by
 CC Streptococcus bacteria, particularly *S. agalactiae* and *S. pyogenes*.
 CC Nucleic acids encoding (I) are used to detect Streptococcus in a
 CC biological sample. (I) is used to determine whether a compound binds to
 CC used as a vaccine or diagnostic composition. The disease caused by
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic
 CC acid encoding (I) may be used to recombinantly produce (I) and may be
 CC used in gene therapy. Antibodies to (I) are used for affinity
 CC chromatography, immunoassays, and distinguishing/identifying
 CC Streptococcus proteins
 XX
 XX Sequence 676 AA;
 QY Query Match 28.1%; Score 663.5; DB 5; Length 676;
 Best Local Similarity 35.3%; Pred. No. 1.1e-56;
 Matches 169; Conservative 82; Mismatches 195; Indels 33; Gaps 13;
 QY 1 MAMVFPSELVNGYDVAATMAAGEMPWMSL-FG-LDVAQAGYQGTLPVLVSVMLATIEKF 58
 DB 212 ICLVSPQLLNAYSVASTSAADIKNWSWNGYFTVQKIGYQAQVIPPALLAGLSLSYLEIF 271
 QY 59 LHKLKGTADFLLTPVLLTLLTGFTFIATGP-----AMRWGVLDVAHGLQGLYDFGPGVG 114
 DB 272 WRKHIPVSVSMIFVFPFLLSPFAILTAHTVLGPIGWTILGKMWISAIVLGLT-----GPK 325
 QY 115 ---GLLGLVSPVITGLHQSPPELLEFNQGSF---IFATASMANIAQGAACLVAF 168
 DB 326 WLFGAIFGALYAFVITGLHMTNAIDTLADTKHTTGLWPMIALSNIAQGSVAIVAY 385
 QY 169 FLAK-SEKGLAGASGVSAVLGITTEPAIFGVNLRWRPFFIGIGTAAIGALIALFNK 227
 DB 386 FMHRHDEKEAQISLPAISAYLVGVTPEALGVNVKVIYTPFVAGMIGSVAGLLATTENQ 445

This Page Blank (uspto)